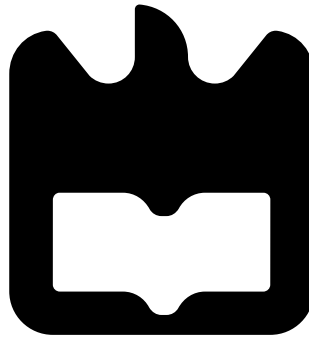




**Joaquim Manuel
de Sousa Fonseca**

Visão computadorizada em condução autónoma





**Joaquim Manuel
de Sousa Fonseca**

Visão computadorizada em condução autónoma

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica dos Professores Dr. Bernardo Cunha e Dr. Artur Pereira, Professores do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

o júri / the jury

presidente / president

Tomás Oliveira e Silva

Professor Associado da Universidade de Aveiro

vogais / examiners committee

Agostinho Gil Teixeira Lopes

Investigador Auxiliar da Universidade do Minho

Manuel Bernardo Salvador Cunha

Professor Auxiliar da Universidade de Aveiro

Artur José Carneiro Pereira

Professor Auxiliar da Universidade de Aveiro

**agradecimentos /
acknowledgements**

Ao meu orientador, Professor Bernardo Cunha, e ao meu co-orientador, Professor Artur Pereira, por todo o tempo e apoio disponibilizado ao longo da Dissertação de Mestrado.

A todas as pessoas envolvidas, até à data, nos projectos Icaro e Rota, por tudo o que desenvolveram e ajudaram a construir.

A todos os meus amigos próximos que me ajudaram a evoluir ao longo de todo o percurso universitário, tanto a nível profissional como a nível pessoal.

À Sara, por toda a força, carinho, amizade e paciência ilimitada.

Aos meus pais por me terem apoiado ao longo de todo o meu percurso académico.

Resumo

O Icaro é um veículo desenvolvido na Universidade de Aveiro, resultante de um trabalho pluridisciplinar que integra várias áreas tais como informática, electrónica e mecânica. O desenvolvimento deste veículo iniciou-se em 1997, com o objectivo de participação na prova internacional Shell Eco-Marathon.

Neste momento, está em desenvolvimento uma nova plataforma, de nome Hammershark, com o objectivo de criar um veículo inovador, tanto em termos de construção como de capacidades finais, de forma a superar o Icaro. Uma das características mais inovadoras que se espera alcançar é a capacidade de condução autónoma, tendo como apoio a visão computadorizada e outros métodos de percepção sensorial.

Esta dissertação aborda a temática da visão computadorizada para a condução autónoma, sendo um trabalho que descreve de forma geral as técnicas normalmente utilizadas para atingir este fim. Assim, o trabalho realizado não se destina especificamente a ser utilizado nas provas da Shell Eco-Marathon, mas sim em todo o tipo de ambientes. O objectivo é elaborar um estudo exploratório e de síntese de conhecimentos sobre a temática, propondo soluções sempre que possível.

As tarefas principais de um sistema que utilize visão computadorizada para condução autónoma podem dividir-se em duas categorias principais: detecção dos limites laterais da faixa de rodagem e detecção de obstáculos. A estas duas tarefas junta-se também uma auxiliar que se resume pela identificação da zona que define a própria da faixa de rodagem, de forma a facilitar a detecção dos seus limites e dos obstáculos que lá circulem.

O trabalho realizado divide-se então nestas 3 categorias, onde são estudadas várias metodologias, propondo possíveis soluções que consigam alcançar os objectivos, quando possível.

Abstract

Icaro is a vehicle developed at the University of Aveiro, resulting from a multi-disciplinary work that integrates ranging areas such as informatics, electronics and mechanics. The development of this vehicle started in 1997 with the objective of participation in the international race Shell Eco-Marathon.

At this moment, a new platform is being developed, Hammershark, in order to create an innovator vehicle, both in terms of building techniques and final capabilities, in order to overcome Icaro. One of the most innovating features that it is expected to achieve is autonomous driving, with the support of computer vision among other methods of sensory perception.

This thesis addresses the topic of computer vision for autonomous driving, describing in general the techniques commonly used to achieve this goal. Thus, the work is not intended specifically for use in the Shell Eco-Marathon, but in all types of environments. The aim is to develop an exploratory study and synthesis of knowledge on the subject, proposing solutions whenever possible.

The main tasks of an autonomous driving system based on computer vision can be divided into two main categories: detecting track side limits and obstacles. To perform these two tasks is also necessary an auxiliary one, that is identify the area that defines the track in order to facilitate the detection of its boundaries and the obstacles which are located on it.

This work is then divided into these 3 major categories, where various methodologies are studied, proposing, whenever possible, solutions that can achieve the goals.

Conteúdo

Conteúdo	i
Lista de Figuras	iii
1 Introdução	1
1.1 Enquadramento	1
1.2 Objectivos	2
1.3 Motivação	3
1.4 Estrutura	3
2 Condução autónoma	5
2.1 Conceitos básicos	5
2.1.1 Sistemas inteligentes de transporte	5
2.1.2 Sistemas inteligentes de veículos	5
2.1.3 Condução autónoma	6
2.2 Projectos relacionados	6
2.2.1 Eureka Prometheus	7
2.2.2 CMU-Navlab	12
2.2.3 Outros projectos relevantes	15
2.3 Darpa Grand Challenge	16
2.4 Condução autónoma em Portugal	19
2.4.1 Festival Nacional de Robótica	19
2.4.2 Critical-Challenge	20
2.4.3 CyberCar	21
2.4.4 Busca e salvamento	22
3 Visão computadorizada	25
3.1 Aquisição de imagem	26
3.2 Espaços de cor	27
3.2.1 RGB	28
3.2.2 HSV	28
3.3 Pré-processamento da imagem	30
3.3.1 Extracção de contornos	30
3.3.2 Segmentação de cor	31
3.3.3 Operações morfológicas	32
3.4 Análise da imagem	32
3.5 Técnicas de análise de imagem	34

3.5.1	Optical flow	34
3.5.2	Detecção baseada em modelos	36
3.5.3	RALPH: Rapidly Adapting Lateral Position Handler	38
3.5.4	UNSCARF	42
3.5.5	GOLD - System for Generic Obstacle and Lane Detection	44
4	Análise de imagem	49
4.1	Biblioteca OpenCV	50
4.2	Ferramentas de manipulação e percepção de cor no OpenCV	50
4.3	Impacto das sombras	52
4.4	Detecção da faixa de rodagem	57
4.5	Detecção dos limites da faixa de rodagem	65
4.6	Detecção de obstáculos	71
5	Avaliação crítica de resultados	87
5.1	Detecção da faixa de rodagem	87
5.2	Detecção dos limites da faixa de rodagem	89
5.3	Detecção de obstáculos	91
6	Conclusões e trabalho futuro	95
	Bibliografia	97

Lista de Figuras

1.1	O Icaro.	1
1.2	Plataformas robóticas para condução autónoma.	2
1.3	O Hammershark.	2
2.1	VITA.	8
2.2	Modelo conceptual do VITA.	9
2.3	Diagrama representativo do fluxo da informação.	10
2.4	Detecção das faixas laterais.	11
2.5	Detecção de obstáculos.	11
2.6	Detecção de intersecções.	12
2.7	Modelo de um sistema de processamento visual.	12
2.8	Os protótipos Navlab 1 e 11	13
2.9	Exemplo de uma rede neuronal.	14
2.10	Rede neuronal utilizada no ALVINN.	14
2.11	Detecção de obstáculos por rastreio de scanner	15
2.12	Exemplos de estradas: não pavimentada e pavimentada.	17
2.13	Momentos do Darpa Grand Challenge, edição de 2004.	17
2.14	Momentos do Darpa Grand Challenge, edição de 2005.	18
2.15	Momentos do Darpa Grand Challenge, edição de 2007	18
2.16	Os robôs de condução autónoma da Universidade de Aveiro.	20
2.17	Plataforma de um CyberCar da Yamaha.	22
2.18	Os robôs dos projectos ROAZ e FALCOS.	22
3.1	Exemplos de ambientes estruturados e não estruturados.	26
3.2	Sistema de visão omni-direcional.	27
3.3	Sistema de visão monocular.	27
3.4	Sistema de visão estereoscópico.	27
3.5	Espaço de cores RGB.	28
3.6	Divisão de uma imagem em RGB nas suas 3 componentes.	29
3.7	Espaço de cor HSV.	29
3.8	Espaço de cor HSV.	30
3.9	Processo de extracção de contornos.	31
3.10	Exemplo de segmentação.	32
3.11	Operações morfológicas.	33
3.12	Cálculo dos vectores de optical flow.	35
3.13	Utilização de um modelo para detecção de veículos.	36

3.14	Detecção de veículos baseada em modelos.	37
3.15	Características principais de um veículo.	38
3.16	Contornos verticais e horizontais.	38
3.17	Detecção de veículos baseada em modelos.	39
3.18	Aquisição e amostragem de imagem no RALPH.	39
3.19	Determinação da curvatura.	40
3.20	Determinação da curvatura.	41
3.21	Determinação da posição lateral do veículo.	41
3.22	Diagrama de processamento do UNSCARF.	42
3.23	Processo iterativo de classificação dos píxeis.	43
3.24	Resultados em estradas com sombras.	44
3.25	Resultados em estradas com neve.	44
3.26	Sistema de visão estereoscópico.	45
3.27	MOBLAB, veículo autónomo.	45
3.28	Remoção do efeito perspectiva.	46
3.29	Detecção das faixas laterais.	46
3.30	Detecção de um objecto homogéneo ideal.	47
3.31	Detecção de um objecto real.	47
3.32	Cálculo da posição do obstáculo.	48
3.33	Mapeamento da posição calculada do obstáculo na imagem real.	48
4.1	Exemplos de ambientes estruturados e não estruturados.	49
4.2	Representação da componente matiz em OpenCV.	51
4.3	Ferramenta de selecção de pontos.	52
4.4	Segmentação de cor.	53
4.5	Impacto da sombra em ambientes reais.	54
4.6	Recolha de amostras de pontos em zonas de sombra e não sombra.	55
4.7	Gráfico das amostras retiradas das imagens na figura 4.6.	55
4.8	Gráfico relativo às amostras recolhidas nas imagens da figura 4.6.	56
4.9	Segmentação da faixa de rodagem com base em sensores padrão.	58
4.10	Histogramas das regiões de análise.	59
4.11	Histogramas globais das regiões de análise.	59
4.12	Exemplo de má segmentação.	60
4.13	Segmentação da imagem usando sensores padrão.	61
4.14	Histogramas das regiões de análise.	62
4.15	Histogramas das regiões de análise.	62
4.16	Aplicação de operações morfológicas.	63
4.17	Binarização da figura 4.16.	63
4.18	Impacto das sombras no processo de segmentação.	64
4.19	Aplicação da transformada de Hough.	66
4.20	Sensores de linha.	67
4.21	Pesquisa da posição óptima dos sensores.	68
4.22	Ajuste dinâmico da posição dos sensores de linha.	68
4.23	Exemplo de regressões lineares.	69
4.24	Sensores de linha e regressão linear dos pontos encontrados.	70
4.25	Sensores de linha e regressão linear dos pontos encontrados.	70
4.26	Imunidade do processo de detecção dos limites laterais relativamente às sombras.	71

4.27	Exemplo do método de geração e verificação de hipóteses.	72
4.28	Aplicação de um algoritmo de extracção de contornos, Canny.	74
4.29	Aplicação do algoritmo de Sobel para extracção de contornos.	74
4.30	Cálculo dos gradientes vertical e horizontal	75
4.31	Cálculo dos gradientes vertical e horizontal da figura 4.28	76
4.32	Cálculo da simetria das imagem na figura 4.29.	77
4.33	Cálculo da simetria das imagens na figura 4.31.	77
4.34	Detecção de sombras.	79
4.35	Segmentação de cor para encontrar as luzes dos veículos	79
4.36	Segmentação de cor para encontrar as luzes dos veículos	80
4.37	Extracção de cantos.	80
4.38	Extracção de formas rectangulares	81
4.39	Imagens sobre as quais irá ser aplicado o processo de detecção de veículos.	82
4.40	Extracção de características na imagem para detecção de obstáculos.	83
4.41	Geração de hipóteses da localização de veículos.	84
4.42	Verificação das hipóteses da localização de veículos.	85
5.1	Processo de detecção dos limites laterais da faixa de rodagem.	88
5.2	Exemplo de mudança brusca no padrão da estrada.	88
5.3	Impacto das sombras no processo de detecção da faixa de rodagem.	88
5.4	Impacto do aparecimento de veículos no processo de segmentação.	89
5.5	Processo de detecção dos limites laterais da faixa de rodagem.	90
5.6	Impacto do aparecimento de veículos na detecção dos limites laterais.	90
5.7	Impacto da sombra na detecção dos limites laterais da faixa.	90
5.8	Diagrama exemplificativo do processo de detecção de veículos.	92
5.9	Processo de detecção de veículos.	93
5.10	Processo de detecção de veículos.	93
5.11	Detecções erradas de veículos.	94
6.1	Exemplo de mau posicionamento das regiões de análise.	95

Capítulo 1

Introdução

1.1 Enquadramento

O Icaro (ver figura 1.1) é um veículo desenvolvido pelo DETI, Departamento de Electrónica, Telecomunicações e Informática, na Universidade de Aveiro, resultante de um trabalho pluridisciplinar que integra várias áreas tais como informática, electrónica e mecânica [1]. O desenvolvimento deste veículo iniciou-se em 1997, com o objectivo de participação na prova internacional Shell Eco-Marathon [2].



Figura 1.1: O Icaro.

O trabalho desenvolvido ao longo dos anos, permitiu incorporar no veículo um conjunto de tecnologias por forma a possibilitar o controlo remoto do mesmo, bem como a sua monitorização tanto para informar o piloto como a equipa que se encontra nas boxes, recorrendo neste caso a telemetria. Os componentes do veículo encontram-se ligados entre si através de uma arquitectura distribuída utilizando uma rede CAN, por forma a ser possível actuar sobre estes mesmos componentes de forma electrónica através de um módulo de controlo. O interface entre o piloto e o veículo é feito através do volante que envia para o módulo de controlo as acções executadas pelo piloto, sendo possível desta forma ligar e desligar o motor, buzinar e actuar sobre o acelerador. Esta ligação entre o volante e o módulo de controlo pode ou não recorrer a um PDA que funciona como gateway, ligado por bluetooth, para estabelecer esta ligação [3].

Desta forma criou-se a possibilidade de conduzir o veículo dentro ou fora do mesmo. Para o modo de condução remota é possível utilizar o volante fora do veículo, desde que a distância permita a ligação bluetooth, ou pode também conduzir-se o veículo das boxes através de telemetria [3].

Este veículo revelou-se essencial na aquisição de conhecimentos e experiência na área da condução remota. O Icaro não é o único veículo desenvolvido no DETI, tendo sido já criados outros

projectos embora com diferentes finalidades. Os projectos Charrua e Rota (ver figura 1.2), entre outros, foram criados com o objectivo de efectuar condução autónoma em ambientes internos estruturados. A construção destes veículos permitiu adquirir um nível bastante elevado de conhecimento nas técnicas de condução autónoma, utilizando já sensores de percepção visual.

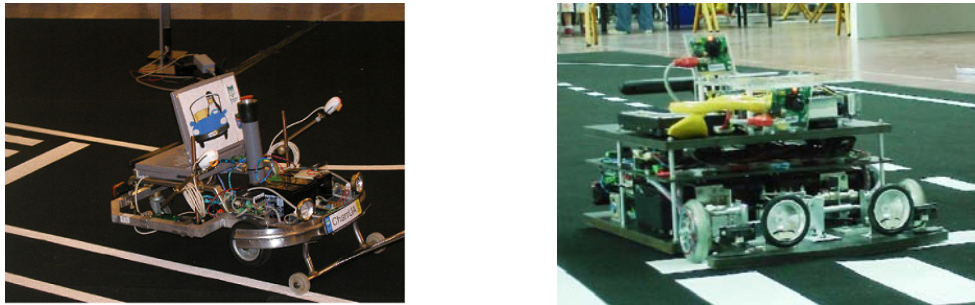


Figura 1.2: Plataformas robóticas para condução autónoma. O Charrua à esquerda e o Rota à direita.

Desta forma, conjugando as experiências e conhecimentos adquiridos através do desenvolvimento do Icaro, Charrua e Rota, entre outros, encontra-se neste momento a ser desenvolvido um novo veículo, o Hammershark (ver figura 1.3), que já se encontra em fase de conclusão. Este veículo tem como principal objectivo a participação nas provas da Shell Eco-Marathon, permitindo desta vez a sua condução autónoma.

O Hammershark tem como objectivo fazer evoluir o conhecimento obtido em condução autónoma de ambientes interiores para ambientes exteriores, tendo em conta todos os novos desafios que esta tarefa implica. Os ambientes exteriores são bastante mais complexos que os interiores, devido à variedade de elementos e interacções imprevisíveis entre eles, o que torna este objectivo bastante difícil de alcançar.

O conhecimento e experiência em condução remota, obtidos no desenvolvimento do Icaro, irão ser também aplicados neste projecto, uma vez que já no caso do Icaro os componentes eram actuados de forma electrónica, tal como é necessário efectuar no Hammershark.



Figura 1.3: O Hammershark.

1.2 Objectivos

Um dos factores que mais influencia a eficiência dos sistemas de condução autónoma é a escolha dos mecanismos de percepção sensorial a utilizar. Estes mecanismos permitem assim a um sistema de condução autónoma perceber o mundo que o rodeia tomando de seguida as decisões adequadas. Existe

um vasto número de equipamentos sensoriais possíveis de usar num sistema deste tipo, dependendo das necessidades e finalidade a que o veículo se destina e do orçamento disponível.

Este trabalho tem como objectivo abordar genericamente vários sistemas de condução autónoma já desenvolvidos, e de forma mais específica os mecanismos de percepção visual e métodos de análise de imagem. Desta forma será elaborado um estudo e recolha de conhecimento sobre toda a temática relacionada, por forma a perceber qual a possibilidade de integrar este tipo de mecanismos no Hammershark.

Será feita uma abordagem aos setups possíveis para um sistema de aquisição de imagem, bem como uma descrição de várias técnicas de processamento e análise de imagem consideradas relevantes, por forma a abranger abordagens tão distintas quanto possível.

As características das estradas onde será possível o veículo conduzir não são especificadas, podendo tratar-se de uma prova com marcas na estrada visivelmente identificadas, ou apenas com relva ou outro tipo de delimitadores que não são previamente conhecidos. Assim, o trabalho realizado não se destina especificamente a ser utilizado nas provas da Shell Eco-Marathon, mas sim em todo o tipo de ambientes.

Um sistema que utilize visão computadorizada para condução autónoma tem como tarefas principais a detecção dos limites laterais das faixas, bem como a detecção dos veículos que lá circulem. Existem também outras tarefas auxiliares de pré-processamento, por forma a extrair ou realçar certas características na imagem para facilitar as tarefas seguintes, como por exemplo a detecção da zona da faixa de rodagem.

O trabalho será realizado então por forma a disponibilizar um estudo sobre as várias abordagens possíveis, em cada tarefa, propondo sempre que possível uma solução para o problema.

1.3 Motivação

Até à presente data os trabalhos realizados no DETI em termos de condução autónoma centraram-se sempre no uso interior, fazendo com que este trabalho seja importante no avanço do conhecimento das técnicas e tecnologias para condução autónoma no exterior. A condução autónoma em ambiente exterior implica novas tarefas, pelo que o estudo e as soluções aqui referidas têm como objectivo trazer algo de novo, o que é por si só uma motivação extra.

O objectivo deste trabalho, tal como já foi referido anteriormente, é efectuar um estudo mais específico sobre os métodos de aquisição e análise de informação visual. Mesmo que o trabalho realizado não seja especificamente desenvolvido a pensar apenas nas provas da Shell Eco-Marathon, é expectável que contribua para desenvolver uma solução a implementar no Hammershark. Esta é também uma motivação extra, uma vez que a prova internacional Shell Eco-Marathon é bastante importante na sua área, sendo o Hammershark um projecto de referência no DETI.

1.4 Estrutura

Esta dissertação divide-se em 6 capítulos. O capítulo 1 é uma abordagem geral à temática estudada nesta dissertação bem como o seu enquadramento, definição dos seus objectivos e motivação.

O capítulo 2 tem como objectivo a introdução à condução autónoma, definindo alguns conceitos básicos e algum do trabalho mais importante realizado até hoje.

O capítulo 3 é uma abordagem à temática da visão computadorizada, descrevendo os tipos de sistemas de aquisição de imagem bem como as tarefas consideradas mais relevantes em sistemas que utilizem esta metodologia. Este capítulo contém ainda uma descrição de várias técnicas de análise

de imagem consideradas relevantes no panorama da visão computadorizada, abrangendo tanto quanto possível abordagens distintas por forma a disponibilizar uma visão genérica de várias metodologias.

O capítulo 4 é uma descrição de todo o trabalho realizado, dividido por categorias, para melhor compreensão das abordagens feitas a cada fase de desenvolvimento, bem como as dificuldades e soluções encontradas.

O capítulo 5 é uma discussão crítica dos resultados obtidos através da realização do trabalho descrito no capítulo anterior.

O capítulo 6 é a parte final deste trabalho, resumindo-se a uma conclusão e possível trabalho futuro.

Capítulo 2

Condução autónoma

2.1 Conceitos básicos

2.1.1 Sistemas inteligentes de transporte

O crescimento exponencial de veículos nas últimas décadas levou a que hoje em dia as infra-estruturas de transporte estejam saturadas. Esta situação afecta a vida das pessoas no seu dia-a-dia, criando atrasos nos transportes, congestionamentos, acidentes e uma elevada poluição. Numa época em que as exigências das pessoas são chegar cada vez mais rápido aos seus destinos e as preocupações ambientais são crescentes, o melhoramento das condições das redes de transporte e sua optimização são alvo de cada vez mais investigação e desenvolvimento.

Desde que esta situação começou a ser entendida como um problema real, foi criado o conceito de ITS (Intelligent Transport Systems) como forma de tentar resolver, ou pelo menos minimizar, alguns dos problemas mencionados anteriormente, assim como trazer melhorias nos anos futuros. Este novo conceito pode ser definido pelo esforço de adicionar capacidades tecnológicas a veículos e infra-estruturas de transportes, por forma a melhorar a eficiência dos mesmos, seja em termos de tempo de deslocação, consumos, segurança ou efeitos para o ambiente. O desenvolvimento destas novas tecnologias permitirão assim tornar a interacção veículo-estrada mais segura e mais eficaz [4, 5].

Muitas são as tecnologias que podem servir de apoio aos sistemas inteligentes de transporte. Algumas estão a ser alvo de mais investigação e desenvolvimento por serem consideradas mais promissoras e capazes de melhores resultados. Entre estas tecnologias estão por exemplo: a visão computadorizada, os radares, referenciação magnética, mapas digitais e GPS [6].

2.1.2 Sistemas inteligentes de veículos

Como uma sub-área das ITS, os IVS (Intelligent Vehicle Systems) representam a tecnologia que é incorporada nos veículos, assistindo ou controlando certas acções de condução. O objectivo é dotar os veículos de sensores, actuadores, métodos de processamento de informação e inteligência artificial, para que estes sejam capazes de perceber o ambiente e actuar correctamente em cada situação.

A interacção destes sistemas computadorizados com o condutor pode ir desde a condução completamente autónoma à simples assistência na condução como forma de aviso ou aconselhamento. Exemplos de sistemas que pertencem a esta última categoria são: visão nocturna, cruise-control adaptativo, aviso/prevenção de colisões [6].

2.1.3 Condução autónoma

A necessidade de por vezes aceder a ambientes que possam ser muito hostis para o ser humano, assim como a necessidade de resolução dos problemas referidos anteriormente causados pelo excesso de tráfego, despoletaram no mundo um grande interesse no desenvolvimento de uma nova área de investigação: condução autónoma de veículos. Este interesse de desenvolvimento de tecnologias inteligentes nasceu já há cerca de 20 anos atrás. As instituições governamentais foram inicialmente quem mostrou interesse no desenvolvimento destas capacidades, quer para uso militar ou apenas para controlo de tráfego.

Um sistema autónomo pode ser definido como algo que utiliza vários órgãos sensoriais para perceber o mundo que o rodeia, para construir uma ponte entre o mundo exterior e a sua representação interna do mesmo, com o objectivo de tomar autonomamente as decisões mais correctas, tal como qualquer condutor humano. Pode-se falar de uma condução autónoma total ou apenas parcial, dependendo do objectivo para o qual o sistema se propõe.

Se referirmos sistemas que estejam adaptados para o uso civil em condições de condução em estrada, podemos então falar de vários tipos de capacidades que podem ser desenvolvidas, tais como detecção e seguimento dos limites laterais da estrada, detecção de obstáculos, manutenção da distância de segurança, planeamento de trajectória bem como ajudas em manobras como estacionamento ou ultrapassagem.

Todas estas capacidades dependem maioritariamente dos sistemas de percepção utilizados, e da quantidade/qualidade de informação que deles é possível retirar. É nos meios de percepção que reside a maior instabilidade de um sistema deste tipo, uma vez que não é possível que os meios de percepção sejam completamente eficazes na informação que disponibilizam, devido à heterogeneidade de ambientes em que é possível os mesmos circular. Numa situação óptima, se a fusão de todos os órgãos sensoriais disponíveis num veículo fosse capaz de em qualquer situação devolver informação completamente precisa e completa sobre o ambiente que rodeia o veículo, seria então muito mais simples conseguir construir sistemas com mais fiabilidade que se pudessem aproximar às capacidades humanas de condução. Até ser atingido este nível de perfeição, a condução completamente autónoma de veículos sem qualquer interacção humana é apenas uma miragem [7, 8].

De uma forma muito resumida podemos definir um sistema de condução autónoma pelas seguintes características:

1. Capacidade de perceber o ambiente que o rodeia.
2. Capacidade de perceber a sua localização neste mesmo ambiente.
3. Capacidade de planear a sua trajectória.
4. Capacidade de actuar nos órgãos mecânicos do veículo para efectuar as acções planeadas.

2.2 Projectos relacionados

Após alguns conceitos introdutórios, irão ser referidos nesta secção alguns dos projectos considerados com mais relevância na história da condução autónoma. Este projectos iniciais tiveram uma importância fulcral no aparecimento e desenvolvimento das primeiras técnicas e tecnologias associadas à condução autónoma.

O projecto Prometheus será o primeiro projecto a ser referido, uma vez que é um dos projectos pioneiros, resultado de um consórcio de unidades de ensino, investigação e empresas, com resultados

afirmados nesta área. O projecto CMU-Navlab pela sua dimensão e quantidade de veículos construídos ao abrigo do mesmo, irá ser referido logo de seguida. Este foi um projecto que, tal como o Prometheus, apareceu quando ainda muito pouco se tinha feito no campo da condução autónoma.

Nesta secção irão ser referidos com algum pormenor os sistemas que foram desenvolvidos ao abrigo destes projectos, bem como os resultados alcançados pelos mesmos.

2.2.1 Eureka Prometheus

Na Europa, o projecto PROMETHEUS (PROgramM for a European Traffic with Highest Efficiency and Unprecedented Safety) teve início em 1986, sendo um projecto pioneiro que iria fomentar o aparecimento de mais grupos de investigação. Este projecto, criado na universidade de Munique, estava ao cargo de uma equipa liderada por Ernst Dickmanns.

Ao longo do seu desenvolvimento, estiveram envolvidos mais de 13 construtores de veículos, várias unidades governamentais e ainda universidades de 19 países Europeus. Graças a este projecto, várias técnicas de condução autónoma foram concebidas e demonstradas [7].

Protótipos desenvolvidos

Como um dos principais resultados desta investigação surgiu nos anos 80 um protótipo, **VaMoRs** (também referido como *vamors-I*), baseado numa carrinha Mercedes de 5 toneladas que foi desenvolvido na universidade de Munique. Este protótipo demonstrou-se capaz de reconhecer a curvatura horizontal e vertical da estrada, a sua largura, número de faixas, a sua posição e orientação em relação à mesma e ainda a detecção de objectos [9].

Quanto ao hardware utilizado, o veículo era composto por duas câmaras frontais para recolher informação visual, que depois seria processada por 23 transputers¹ organizados em módulos. Com esta arquitectura, e tendo em conta as limitações computacionais da altura, foi conseguido que os exigentes requisitos computacionais fossem satisfeitos. Os módulos executavam várias tarefas distintas tais como análise da imagem, modelação de objectos, integração dos dados sensoriais e controlo [10].

As capacidades deste protótipo foram demonstradas em 1985, numa auto-estrada Alemã, onde o mesmo foi capaz de conduzir autonomamente com velocidade de cerca de 100km/h. Para informação mais detalhada ler [9].

Em 1990, com hardware mais reduzido mas com superior poder computacional, surgiu um novo protótipo, o **VAMP** (também referido como *vamors-p*), desta vez um veículo de passageiros. Este protótipo mostrou-se capaz de conduzir autonomamente em tráfego normal, já com capacidade de ultrapassar veículos mais lentos, atingindo a velocidade máxima de 180km/h. Em 1995, conseguiu completar um circuito de 158km numa auto-estrada Dinamarquesa, sem qualquer intervenção humana [11].

Quanto ao hardware, o protótipo era constituído por 2 conjuntos de câmaras, para permitir ter uma perspectiva para a frente e para a retaguarda do veículo, sendo possível detectar e seguir até 5 obstáculos, em cada direcção.

Cada conjunto de câmaras continha no mínimo duas câmaras CCD, por forma a ser possível explorar os modos bifocal ou multi-focal. Diferentes configurações das câmaras permitiam adaptar o veículo para diferentes modos de actuação, conseguindo atingir uma maior distância de visualização, até 100m, para ser possível alcançar velocidades superiores a 130km/h. No entanto, um tempo de

¹micro-processadores concebidos para processamento paralelo

resposta de 40ms permitia ao sistema actuar perfeitamente em tempo real, a curtas distâncias, conseguindo tomar acções muito rapidamente [11].

Em termos de suporte computacional, este protótipo era constituído por 60 transputers, incluindo também um computador que servia de host a todos estes micro-processadores ao contrário da arquitectura anterior, o **VaMoRs** [12].

Em 1991, como resultado do trabalho efectuado no âmbito deste mesmo projecto, EUREKA-PROMETHEUS, foi apresentado em Torino mais um protótipo: **VITA** (VIsion Technology Application). Este veículo teve a participação complementar da BOSCH e da HELLA, para construção e desenvolvimento das partes de electrónica e componentes automóveis mais complexos.

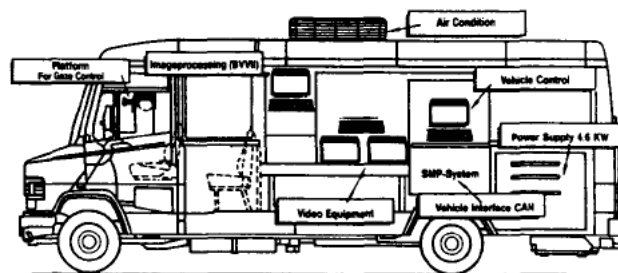


Figura 2.1: VITA [5].

Este veículo estava equipado com uma arquitectura um pouco mais complexa que as anteriores, utilizando uma estrutura CAN² para transmitir a informação sensorial e os sinais de actuação entre todos os controladores envolvidos. A plataforma de trabalho foi desenvolvida numa carrinha Mercedes-Benz, com o intuito de oferecer espaço e condições de energia e temperatura propícias de trabalho. O resultado foi um laboratório móvel de testes, que possibilitava aos investigadores instalarem-se no seu interior para testarem as suas aplicações.

O âmbito deste novo protótipo foi envolver vários institutos de investigação com o objectivo de demonstrar algumas capacidades inovadoras no campo da visão computadorizada. Entre estes institutos estiveram envolvidos a Universidade das Forças Armadas em Munique, o Instituto Fraunhofer em Karlsruhe e as Universidades de Bochum e Paderborn [5].

Este veículo demonstrou-se capaz de algumas tarefas inovadoras de condução autónoma tais como:

1. Seguimento da estrada com bastante imunidade a sombras
2. Regulação automática de velocidade consoante a curvatura da estrada.
3. Detecção de obstáculos na faixa de rodagem.
4. Sistema stop&go, sistema que permite seguir um veículo mantendo uma distância fixa pré-determinada.

²Controller Area Network: um barramento série de microcontroladores de rede que conecta dispositivos, sensores e actuadores de um sistema ou sub-sistema.

5. Detecção de intersecções (zonas onde a estrada em que o veículo circula é interceptada por outras estradas).
6. Capacidade de condução em condições adversas, como é o caso de más condições climatéricas.

O hardware de baixo nível do veículo era composto por 22 controladores CAN para suportar a transferência de informação sensorial e de ordens de actuação entre todos os componentes. Esta arquitectura permitiu assim criar uma camada de abstracção a este nível, disponibilizando uma interface de comunicação com todos os componentes para que a comunicação com estes se tornasse mais transparente.

A aplicação de controlo residia num computador, chamado co-piloto, que comunicava com cada controlador por esta interface de comunicação. Para cada tarefa simples tal como travar, acelerar ou mudar de direcção existe um controlador dedicado, que pode ser controlado directamente pelo co-piloto ou por controladores intermediários. Um destes intermediários é o SULOC, controlador que supervisiona a aceleração e a travagem de acordo com a distância longitudinal (ver figura 2.2).

Por fim, existe ainda um interface que permite a um humano entrar também no sistema (*human in the loop*), e interagir com a velocidade e direcção. Desta forma o sistema pode controlar o veículo para que este circule de forma autónoma, ou pode apenas funcionar como sistema de aviso.

Esta metodologia tornou o veículo uma boa plataforma de testes.

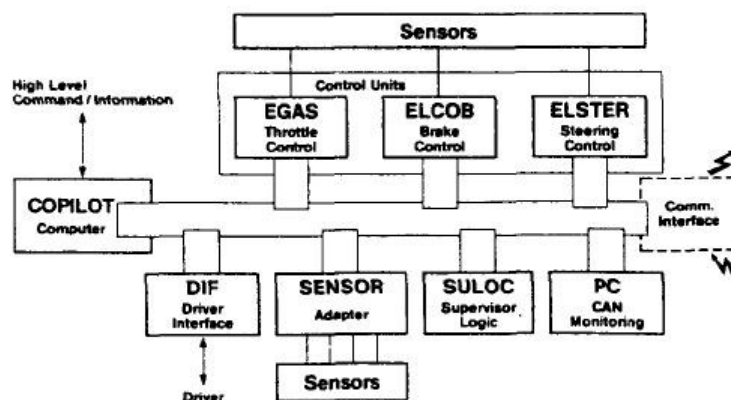


Figura 2.2: Modelo conceptual do VITA [5].

A aquisição de imagens era feita por duas câmaras monocromáticas CCD, com diferentes distâncias focais. A câmara com menor distância focal era usada para detecção da faixa lateral, enquanto que a câmara com maior distância focal era usada para detecção de obstáculos a média/longa distância.

O dispositivo de aquisição permitia ainda a movimentação das câmaras, para melhor aquisição em situações de curvas com maior ângulo de curvatura. Este dispositivo transferia informação a cerca de 50Mbits/s, o que superava as capacidades dos links físicos entre transputers (aproximadamente 20Mbits/s). Para resolver este problema foram implementados links paralelos e utilizados métodos de análise parcial da imagem, apenas em zonas de interesse definidas, para diminuir a quantidade de informação que era obtida pelas câmaras.

Sistemas de controlo desenvolvidos

Um dos grandes problemas de análise de sequências de imagens em tempo real é o seu elevado custo computacional e, como consequência, um elevado tempo de execução que pode por vezes tornar

o sistema ineficaz. Para minimizar este problema, a equipa de Ernst Dickmanns desenvolveu um sistema que usava uma estratégia baseada em métodos de predição para permitir a análise de partes específicas das imagens, rejeitando outras irrelevantes.

O método consistia em construir modelos 4D internos para representar a estrada e os veículos, assim como a própria posição e orientação, tendo em conta o espaço e o tempo. Estes modelos eram recursivamente actualizados e corrigidos de acordo com a informação sensorial e com a informação de imagens anteriores, utilizando filtros de Kalman. Este algoritmo exige inicialmente a necessidade de uma pesquisa mais extensiva para determinar os parâmetros da estrada, para poder formar o modelo inicial (ver figura 2.3). Após esta tarefa inicial, torna-se muito mais fácil manter o modelo actualizado, explorando a condições de continuidade temporal.

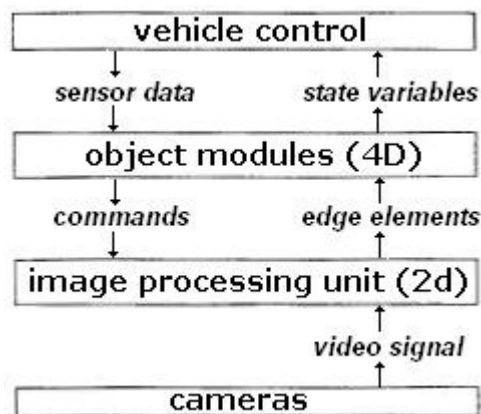


Figura 2.3: Diagrama representativo do fluxo da informação [10].

Com estes modelos espacio-temporais foi conseguida alguma imunidade a ruídos e incertezas, resultando assim num modelo que conseguiu alcançar alguma robustez para efectuar condução autónoma. Este modelo é no entanto adverso a mudanças radicais na estrutura da estrada, uma vez que o modelo necessita de algum tempo para se adaptar. Para informação mais detalhada sobre este sistema, ler [13, 14]

O algoritmo de detecção e seguimento dos limites laterais da faixa de rodagem era baseado na análise de 8 áreas de interesse, 4 em cada lado da imagem, que tinham como objectivo estarem centradas nestes próprios limites (ver figura 2.4). Após conseguir inicialmente localizar estas áreas de interesse, estas eram reposicionadas em cada nova imagem, tendo em conta a sua posição anterior e as possíveis mudanças na posição da faixa na imagem actual. Este método permitia assim obter uma aproximação da posição dos limites laterais, assim como a posição e orientação do veículo em relação aos mesmos [5]. A cada nova imagem capturada, a plataforma de aquisição de imagem reposicionava a câmara por forma a que o centro de gravidade destas 8 áreas de interesse estivesse posicionado no centro da imagem [5].

O algoritmo de detecção de obstáculos permite um controlo longitudinal do veículo a cerca de 100m à sua frente, dependendo da sua velocidade. A estratégia deste algoritmo é baseada numa metodologia de criação e verificação de hipóteses. Inicialmente são procuradas zonas de imagem que possam ser candidatas a serem consideradas como veículos (ver figura 2.5). Posteriormente é feita uma verificação destas mesmas hipóteses, de acordo com alguns critérios baseados em simetria, forma e localização.

Uma característica fundamental de praticamente todos os veículos é a relação simétrica da sua retaguarda, em relação a um eixo vertical. É principalmente nesta relação que o algoritmo se baseia



Figura 2.4: Detecção das faixas laterais [5].

para fazer a validação dos candidatos a veículos. Se os critérios forem satisfeitos, estes veículos são seguidos e é calculada a velocidade relativa e distância em relação ao veículo.

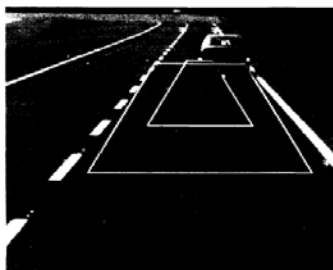


Figura 2.5: Detecção de obstáculos [5].

A partir deste sistema, é possível parar ou reduzir velocidade para evitar colisões, ou ainda aumentar a velocidade para manter uma distância constante para o veículo que se encontra à sua frente. Quanto às ultrapassagens, ou mudanças de faixa, é necessária confirmação por parte do condutor, uma vez que o sistema não possui uma câmara à retaguarda para efectuar estas operações com segurança.

Um outro módulo permitia a detecção de intersecções analisando a configuração das faixas laterais (ver figura 2.6). Em termos de hardware, este módulo era composto por vários transputers dedicados a este mesmo processamento.

O seu método de funcionamento era baseado em modelos, ou seja, vários modelos de intersecções eram carregados para o algoritmo para poderem depois ser comparados com uma possível intersecção. Numa situação real em que o veículo se aproximava de uma intersecção, vários modelos de intersecções eram comparados. Ao encontrar um modelo semelhante à intersecção que se aproxima, o algoritmo pode tomar várias acções, entre diminuir a velocidade, ou examinar com mais cuidado a mesma intersecção [5].

A camada de controlo é implementada uma vez mais num sistema de transputers dedicados, onde é feito todo o planeamento assim como as respectivas acções. Neste nível, são formuladas estratégias de controlo de acordo com a informação sensorial recebida. O controlo longitudinal e lateral do veículo é inteiramente realizado a este nível, assim como outras operações tais como a mudança de faixa. Existe ainda um nível superior, que permite a interacção de um condutor, introduzindo mais um elemento no sistema.

Tal como é ilustrado na figura 2.7, este sistema é composto por vários sub-módulos (reconhecimento das faixas laterais, obstáculos e intersecções) que adquirem informação do ambiente, e constroem um modelo dinâmico do mundo, ou da informação relevante do mesmo.

Os módulos de planeamento e decisão, tendo como informação base este mesmo modelo formado pelos órgãos sensoriais, são capazes de derivar uma acção que permita ultrapassar com um grau



Figura 2.6: Detecção de intersecções [5].

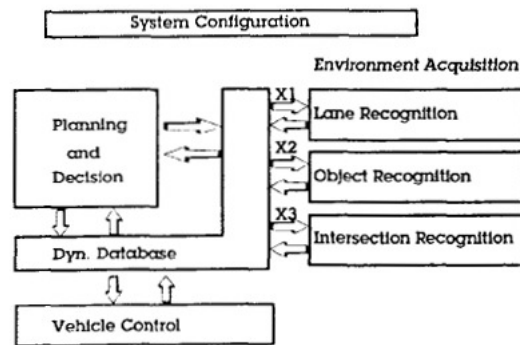


Figura 2.7: Modelo de configuração do sistema de processamento visual [5].

confiável de segurança uma qualquer situação em tempo real.

2.2.2 CMU-Navlab

O grupo NAVLAB foi fundado na universidade de Carnegie Mellon, em Pittsburgh, com o objectivo de desenvolver veículos com capacidades de condução autónoma. Criado em 1984, este grupo conseguiu desenvolver um total de 11 veículos, desde o Navlab 1 até ao Navlab 11.

O objectivo desta investigação era a criação e desenvolvimento de aplicações capazes de conduzir, ou assistir a condução, em ambientes reais, sem qualquer necessidade de alteração de infra-estruturas ou intercomunicação entre veículos. O grupo Navlab nasceu ao abrigo de um outro projecto, com o intuito de reduzir a presença humana em ambientes hostis, tal como é o caso de operações militares. No entanto, as técnicas desenvolvidas são também adaptáveis aos veículos civis, uma vez que podem ser usadas para controlar o veículo ou apenas para avisar os condutores de condições perigosas [15, 16].

Protótipos desenvolvidos

Os veículos construídos por este grupo de investigação eram compostos por uma fusão de órgãos sensoriais, que incluíam sensores laser, câmaras de vídeo, radares, sensores de infravermelhos, termográficos e sensores de inércia. Alguns tinham espaço suficiente para transportar os investigadores, para que estes pudessem verificar “on-the-run” os resultados obtidos. Estes veículos eram também compostos por mecanismos automáticos de controlo da aceleração e direcção do veículo, que podiam

ser actuados pelos sistemas autónomos de condução. Estas características tornavam os veículos em plataformas perfeitas para testar os sistemas desenvolvidos.

Inicialmente o primeiro veículo desenvolvido, Navlab1, era um perfeito laboratório móvel, sendo dotado de toda a tecnologia referida anteriormente com uma grande capacidade electrónica e computacional. Tinha espaço para 4 pessoas na parte de trás, o que permitia uma maior proximidade entre os investigadores e a própria experiência, assim como eliminava a necessidade de transferência de informação para análise [15, 16, 17].

Com o desenvolvimento de novos equipamentos mais leves e menos dispendiosos em termos de gastos energéticos, foi possível alcançar uma nova mobilidade e portabilidade que permitiu que estes sistemas fossem implementados em carros comuns de passageiros. O último veículo desenvolvido no âmbito deste projecto é prova disto mesmo, sendo um veículo de passageiros comum, de dimensões regulares (ver figura 2.8).



Figura 2.8: Os protótipos Navlab 1 e 11

Sistemas de controlo desenvolvidos

Em 1984, juntamente com a Intel, o grupo Navlab iniciou o desenvolvimento de um sistema de processamento denominado de Warp Machine, com o objectivo deste ser capaz de servir as altas exigências computacionais do seu sistema. Este sistema computacional é uma forma particular de processamento paralelo em que as unidades de processamento têm capacidades de armazenamento individual, e estão ligadas entre si por forma a facilitar a rápida comunicação entre elas.

Este sistema de hardware precedeu o desenvolvimento das aplicações de condução autónoma, revelando-se muito importante no sucesso do grupo Navlab. Uma vez que foi desenhado especificamente para estas funções, este sistema potenciou o desenvolvimento de aplicações que tinham inicialmente um elevado custo computacional [18, 19].

A detecção dos limites laterais da faixa de rodagem, assim como o cálculo da distância lateral entre o veículo e estes mesmos limites, é uma característica fundamental em todos os sistemas de controlo autónomo de veículos. O sistema de posicionamento lateral desenvolvido neste projecto é baseado numa arquitectura de simulação de uma rede neuronal de nome ALVINN (Autonomous Land Vehicle In a Neural Network).

Redes neuronais artificiais são sistemas computacionais estruturados numa abordagem baseada em ligações. De uma forma genérica, podem ser descritas por uma quantidade de nós, organizados em n camadas, formando uma rede. Cada nó possui links para outros nós, que comunicam entre si, podendo assim formar um sistema que recebendo x entradas, irá formar y saídas, onde as saídas são função das entradas [16].

Neste caso específico, o ALVINN utiliza como entradas regiões específicas de uma imagem, onde

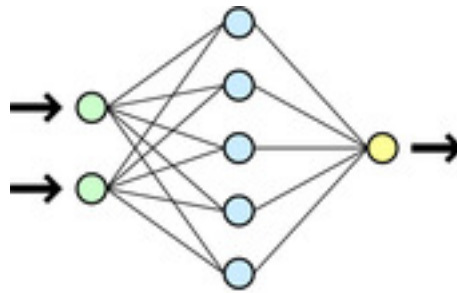


Figura 2.9: Exemplo de uma rede neuronal.

cada região irá alimentar uma entrada desta rede. Desta forma é possível fazer uma análise completa de uma imagem, onde cada região contribui para calcular a saída, que neste caso irá ser a respectiva acção a tomar.

O sistema ALVINN dividia inicialmente a imagem por forma a obter 30x32 regiões, o que iria resultar num total de 960 entradas para este sistema. Estas entradas formam a camada de entrada do sistema. Desta camada são gerados sinais de activação que irão ser transmitidos pela rede até formarem uma camada de saída. Nesta camada estão presentes várias saídas, cada uma representando uma acção. A acção a tomar é escolhida de acordo com a saída que estiver mais activa.

A cada região, ou grupo de regiões, estavam associados determinados pesos, para que contribuissem de forma diferente para o resultado, de acordo com a sua localização e respectiva importância para determinação da saída.

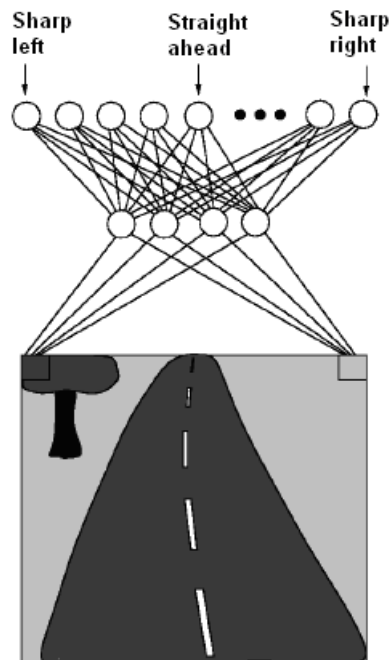


Figura 2.10: Rede neuronal utilizada no ALVINN [16].

A cada nó desta rede está associada uma memória local que permite armazenar alguns dados, tal como informação sobre o passado. Esta característica potencia bastantes capacidades a este sistema, uma vez que desta forma é possível obter um sistema dinâmico capaz de se auto-adaptar de acordo

com novas realidades.

O método comum de funcionamento deste sistema é, tal como já foi anteriormente descrito, a partir de determinadas características na estrada que formam as entradas do sistema, produzir uma acção apropriada que seja resultado de uma ponderação das entradas. Ou seja, é fundamental que os pesos atribuídos a cada região da imagem sejam correctamente definidos para obter um bom funcionamento do sistema. Para definir estes pesos das entradas é possível fazer com que o sistema se auto-configure, aprendendo por exemplo com a condução de um humano.

Ou seja, é possível efectuar um período de configuração em que, com um condutor real, o sistema é capaz de relacionar as acções deste mesmo condutor com as características da estrada em cada momento, formulando ele próprio os pesos que acha correctos para cada entrada do sistema. O sistema ALVINN é portanto capaz de, após um período mínimo de aprendizagem de 3 minutos observando um condutor real, tomar o controlo do veículo continuando a marcha autonomamente. Estas capacidades permitem a este sistema conduzir a velocidades de aproximadamente 90km/h, em vários tipos de estradas [16].

O sistema de detecção e seguimento da estrada desenvolvido no projecto Navlab é usado para fazer um planeamento da trajectória a percorrer pelo veículo. Quando é detectado algum obstáculo que colide com a trajectória planeada para o veículo, o controlo do veículo é passado para o módulo de evasão de obstáculos que altera a trajectória do veículo para conseguir evitar com segurança o obstáculo.

Este sistema era suportado por um sensor laser, com a capacidade de fazer 8 scans por segundo, tendo um alcance de 30 metros com ângulo de 120 graus. Este rastreio era feito consecutivamente, à medida que o veículo se move, tentando criar um mapa com possíveis obstáculos na faixa de rodagem a circular à frente do veículo.

A detecção de obstáculos neste sistema era feita usando uma aproximação poligonal da informação proveniente do sensor laser, que pudesse representar possíveis obstáculos (ver figura 2.11) [20].

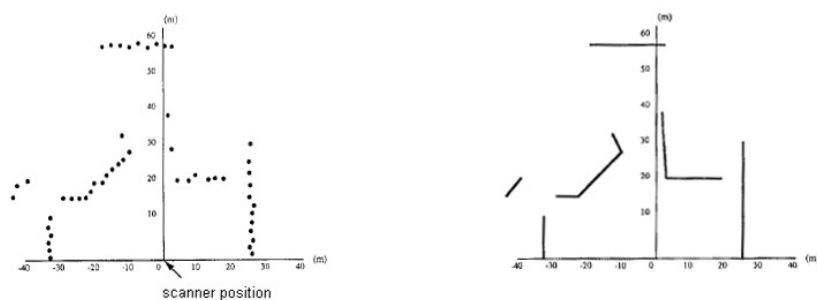


Figura 2.11: Da esquerda para a direita: detecção de obstáculos por rastreio de scanner; aproximação poligonal dos resultados [20].

2.2.3 Outros projectos relevantes

Nos Estados Unidos existiram várias iniciativas como tentativa de resolver os problemas de mobilidade, que envolveram universidades, centros de investigação e construtores automóveis. No fim da década de 1980, resultado de um estudo de um grupo de investigação, foi criado o IVHS America (Intelligent Vehicle Highway Systems), constituído por uma parceria entre entidades privadas e públicas como forma de consolidar os interesses nacionais de desenvolvimento de sistemas inteligentes de

transporte [4]. Este projecto tinha como objectivo criar sistemas que permitissem condução autónoma em auto-estradas.

Após esta fase piloto, em 1995 o governo dos Estados Unidos criou o NAHSC (National Automated Highway System Consortium) e lançou o programa Intelligent Vehicle Initiative em 1997 [21, 22].

No Japão, onde os problemas de mobilidade são ainda mais preocupantes devido à elevada densidade populacional, alguns protótipos de veículos autónomos foram também desenvolvidos ao abrigo de projectos criados, tais como o Intelligent Vehicle, o Personal Vehicle System (PVS) e o Advanced Cruise-Assist Highway System Research Association (AHSRA) [23, 24].

O projecto AHSRA surgiu em 1996 numa iniciativa do Departamento de infra-estruturas e transportes do Japão envolvendo várias entidades entre construtores e centros de investigação, que resultou na criação de várias abordagens para o problema de condução autónoma de veículos. A finalidade da formação deste grupo de investigação era fazer um estudo detalhado sobre os sistemas de transportes, com objectivo de melhoria da segurança, eficiência e conforto. O estudo desde início foi focado na incorporação de sistemas inteligentes tanto nos veículos como nas vias de transporte, com o objectivo de explorar a potencialidade da cooperação veículo-estrada.

Desde 1996 até 1998, foi feito um estudo de viabilidade e definição de requisitos necessários ao desenvolvimento deste tipo de sistemas. Em 2000 e 2002 vários protótipos foram testados em percursos de teste, para melhor perceber quais seriam as possibilidades de integração de alguns sistemas em ambientes reais. Desde então, o trabalho deste grupo de investigação continua a ser perceber os impactos da utilização destes sistemas, na melhoria das condições de transporte [25].

2.3 Darpa Grand Challenge

O Departamento da Defesa dos Estados Unidos criou em 1958 uma agência de investigação, DARPA (Defense Advanced Research Projects Agency) [26, 27], com o objectivo de investigação e desenvolvimento de tecnologia militar. Em 2004, organizada por esta mesma agência, surgiu uma competição para veículos terrestres autónomos. Esta competição tinha como objectivo potenciar globalmente o desenvolvimento de novas tecnologias ligadas à condução autónoma, obrigando construtores automóveis, centros de investigação e universidades a uma fusão de esforços, por forma a criarem e desenvolverem tecnologias capazes de cumprir os objectivos das provas.

A participação nesta prova obrigava as equipas a desenvolverem várias técnicas e equipamentos relacionados com condução autónoma. Era necessária a utilização de equipamentos para recolha de informação externa, tais como GPS, sensores laser, captura de imagem, análise de dados de odometria, para posterior análise e tomada de decisões. Este requisitos criavam assim um compromisso entre os equipamentos a utilizar e o orçamento mais ou menos alto de cada equipa.

A meta imposta pela agência organizadora era de, em 2015, 1/3 dos veículos militares estarem dotados de capacidades de condução autónoma.

Esta competição consiste em provas de longa distância, sem intervenção humana, no menor tempo possível, evitando obstáculos e seguindo um percurso pré-determinado. O percurso da prova pode ser composto por estradas pavimentadas, não pavimentadas ou mesmo zonas de deserto (ver figura 2.12). Os obstáculos podem ser constituídos pelas mais variadas coisas tais como bermas, água, pedras, passagens subterrâneas, equipamentos de construção ou outros.

Era imposta aos concorrentes uma vasta lista de regras, que variavam entre a obrigação de equipar os respectivos veículos com mecanismos de bloqueio em caso de emergência, para prevenir colisões ou outras situações de perigo, até à proibição do uso de qualquer dispositivo de controlo remoto ou

transferência de informação sem fios [27].



Figura 2.12: Exemplos de estradas: não pavimentada e pavimentada.

A primeira prova desta competição realizou-se em 2004, num percurso todo-o-terreno, de aproximadamente 240km, entre os estados da Califórnia e Nevada. Os veículos eram obrigados a cumprir o circuito pré-definido para a prova. Participaram 25 equipas neste evento, das quais nenhuma conseguiu cumprir com sucesso todo o circuito estabelecido, sendo que a máxima distância percorrida por uma equipa foi aproximadamente 12km. O prémio de 1 milhão de dólares, para a equipa que completasse mais rápido o percurso, com um limite de 10h, não foi atribuído.



Figura 2.13: Momentos do Darpa Grand Challenge, edição de 2004.

Em 2005, uma nova edição foi organizada, desta vez com um circuito de aproximadamente 210 km, novamente em todo-o-terreno, no deserto sudoeste dos Estados Unidos. Nesta edição 23 equipas conseguiram passar as provas técnicas iniciais e ganharam o direito de participar no evento principal. Ao contrário da edição anterior, 4 veículos conseguiram realizar todo o percurso da prova, sendo que o mais rápido efectuou o traçado em aproximadamente 7h.

A equipa vencedora, Stanford Racing Team, ganhou 2 milhões de dólares, como recompensa por ser a mais rápida. Esta equipa era maioritariamente formada por alunos da Stanford's School of Engineering. O veículo foi construído sobre a carroçaria original de um Volkswagen Touareg R5, com capacidades todo-o-terreno reforçadas. O veículo era composto por um sistema drive-by-wire gerido electronicamente por 7 computadores, ou seja, todas as actuações “físicas” no veículo, tais como acelerar, travar ou dirigir, eram iniciadas por um sistema electrónico ligado aos componentes mecânicos que desencadeava as acções pretendidas.

Os órgãos sensoriais utilizados para “perceber” o ambiente eram dos mais variados, constituindo um sistema bastante complexo de aquisição composto pelos seguintes sistemas/equipamentos: GPS, odometria, um sensor de inércia, sensores laser, um sistema de radar e dois sistemas de aquisição de imagem, um estereoscópico e outro monocular. Este complexo sistema de aquisição permitia adquirir



Figura 2.14: Momentos do Darpa Grand Challenge, edição de 2005. À direita, o veículo vencedor.

toda a informação com uma frequência entre 10 e 100 HZ, o que permitia ao veículo ser bastante eficaz, mesmo em situações em que era necessário tomar decisões rápidas.

Em 2007, tendo em conta o sucesso das duas provas anteriores, foi novamente realizada uma edição desta prova, sendo a última até hoje. O percurso desta prova teve características mais urbanas, tendo em conta as necessidades actuais de condução autónoma em meios estruturados. Era necessário às equipas a realização de manobras mais técnicas e complexas, tais como estacionamento, ultrapassagem e passagem em intersecções, em ambientes com possível tráfego.

Era obrigatório às equipas realizar um longo e rigoroso período de verificações técnicas e qualificação para a prova principal, que demorou cerca de 8 dias. Das 83 equipas inicialmente inscritas, apenas 35 tiveram direito a participar nas provas de qualificação, das quais apenas 11 conseguiram com sucesso passar as provas de qualificação ganhando direito à participação na prova principal. Como prova do avanço tecnológico alcançado até então, durante as provas de qualificação que foram realizadas em ambiente com tráfego, foi apenas registado um acontecimento de choque entre veículos.

O evento principal era constituído por um percurso com vários checkpoints que cada equipa tinha que percorrer. Esta informação foi disponibilizada apenas 5 minutos antes da prova ter início, criando um verdadeiro desafio de condução autónoma.

No momento do início da prova, 11 veículos autónomos juntamente com 30 não autónomos, iniciaram a sua marcha simulando assim uma situação real de tráfego.



Figura 2.15: Momentos do Darpa Grand Challenge, edição de 2007. À esquerda, vista do traçado; à direita, a equipa vencedora.

A equipa tartan [28], representada na figura 2.15, conseguiu acabar a prova em aproximadamente 4h, seguida da equipa de Stanford, vencedora da edição anterior, que terminou apenas um minuto depois. Este resultado permite demonstrar o avançado nível das equipas a concorrer nesta prova, assim como o potencial da implementação de algumas tecnologias em ambientes reais.

A equipa vencedora era formada por uma aliança entre a Carnegie Mellon University [15], General

Motors, Continental e Catterpillar, potenciada por um sem número de patrocinadores tais como a Google e a Intel, o que demonstra que o sucesso neste tipo de eventos está inevitavelmente ligado ao apoio e disponibilidade financeira de cada equipa. O veículo construído teve como base um Chevy Tahoe [29], composto por mais de doze sensores laser, câmaras e radares para ser capaz de perceber o mundo em seu redor. O seu sistema de controlo, composto por mais de 500 mil linhas de código, permite ao veículo fazer um correcto reconhecimento das características da estrada, bem como obstáculos, para posteriormente planear o seu comportamento e tomar as devidas acções.

Estes resultados demonstram ao mundo que a condução autónoma em ambiente urbano pode um dia tornar-se realidade.

2.4 Condução autónoma em Portugal

Portugal tem uma actividade educativa e de investigação considerável na área de robótica, quando comparado em termos relativos ao número de habitantes e investigadores/professores activos, com outros países de todo o mundo que se dedicam também à investigação nesta área.

Existem já algumas actividades competitivas em Portugal que procuram fomentar a investigação e desenvolvimento no campo da robótica. Nesta área irá ser feita referência ao Festival Nacional de Robótica, que é neste momento a actividade de referência a nível nacional contendo várias modalidades de competição robótica, promovida pela Sociedade Portuguesa de Robótica.

A Sociedade Portuguesa de Robótica foi criada em 2006 como uma associação sem fins lucrativos cujo principal objectivo é o de promover e estimular o ensino, a investigação científica, o desenvolvimento tecnológico e as aplicações (indústria e serviços). Esses objectivos são conseguidos através de várias acções, as quais incluem, entre outras: O Festival Nacional de Robótica, publicações regulares, seminários e encontros [30].

Irá também ser referida uma iniciativa muito recente, organizada pela Critical Move em parceria com o MIT, que se denomina por Critical Challenge.

Na área de projectos de investigação irão ser referidos alguns trabalhos realizados pelo laboratório de sistemas autónomos do Instituto Superior de Engenharia do Porto e pelo laboratório de robótica móvel da Universidade de Coimbra, que são considerados relevantes no panorama nacional em termos de condução autónoma.

O conceito de robótica abrange um sem número de dispositivos e equipamentos, mas nesta secção irão ser referidos alguns trabalhos feitos em Portugal, principalmente aqueles que são desenvolvidos em ambiente académico, apenas na área da condução autónoma uma vez que é a sub-área da robótica que mais se enquadra no tema desta dissertação. Quando possível, irá ser dada mais atenção aos mecanismos de percepção visual destes mesmos projectos.

2.4.1 Festival Nacional de Robótica

O Festival Nacional de Robótica conta já com nove edições, tendo sido iniciado em 2001, com o objectivo de promoção da ciência e tecnologia junto dos jovens do ensino básico, secundário e superior. Esta actividade consiste em competições de robôs de várias categorias, bem como exposição dos mais recentes desenvolvimentos na área da robótica. Decorre anualmente numa cidade distinta, estando a assumir cada vez mais importância no panorama robótico nacional, tendo-se tornado já um evento de referência.

Uma das provas sénior deste festival é a prova de condução autónoma, na qual são estimuladas e testadas as capacidades de conduzir um veículo sem intervenção humana, de acordo com certas

regras. Os veículos têm de cumprir uma pista, que simula um cenário real, tendo que compreender os sinais luminosos mostrados nos ecrãs para indicar qual a manobra a efectuar. A pista normalmente inclui obstáculos, zonas de obras, passadeiras, sinais luminosos e uma zona de estacionamento onde os veículos têm de encontrar um lugar vazio.

A competição decorre num ambiente semi-estruturado. A pista e alguns elementos (passadeira, posição dos sinais luminosos) são conhecidos à priori, mas outros elementos (obstáculos, zona de obras) não. Esta competição testa a capacidade dos veículos conduzirem autonomamente desviando-se de obstáculos, estacionarem e reconhecerem os sinais luminosos [31].

A Universidade de Aveiro concorre normalmente com várias equipas provenientes dos departamentos de Mecânica e de Electrónica Telecomunicações e Informática. O Atlas-MV desenvolvido pelo departamento de Mecânica é o último veículo desenvolvido por este departamento, que conseguiu a vitória no último festival de robótica em 2009. Este veículo além dos sensores de percepção visual para interpretar as linhas laterais e os sinais luminosos, conta também com um laser scanner que permite com grande precisão localizar a posição dos obstáculos que se encontrem na via.

O departamento de Electrónica Telecomunicações e Informática tem também um projecto de um veículo autónomo que participa em todas as edições do FNR, o Rota. Este veículo conta apenas com duas câmaras como órgãos sensoriais, uma voltada para cima para ver o sinal luminoso e outra voltada para o chão para perceber a localização dos limites laterais da faixa (ver figura 2.16).



Figura 2.16: Os robôs de condução autónoma da Universidade de Aveiro. À esquerda, o Rota do Departamento de Electrónica, Telecomunicações e Informática; à direita, o Atlas-MV do Departamento de Mecânica.

O festival nacional de robótica revela-se cada vez mais fundamental para o desenvolvimento de novas técnicas e tecnologias para condução autónoma, funcionando como incentivo para as organizações melhorarem ano após ano.

2.4.2 Critical-Challenge

A Critical Move é um spin off da Critical Software, criada com o objectivo de desenvolver e produzir soluções móveis inovadoras, flexíveis, limpas e seguras na área das tecnologias inteligentes de transportes. Estas tecnologias são criadas com o objectivo de corresponder às necessidades actuais e futuras dos sistemas de transporte [32].

Esta nova empresa encontra-se neste momento a construir um veículo autónomo, o Move, para servir de transporte de pessoas em locais públicos como hospitais, aeroportos, parques industriais ou resorts turísticos. O objectivo deste veículo é servir como complemento dos sistemas de transporte principais, viajando a baixa velocidade e cumprindo distâncias curtas.

A par deste projecto, a Critical Move lançou um desafio em parceria com o MIT-Portugal, a estudantes que queiram participar num concurso para o desenvolvimento de outro protótipo semelhante,

mais avançado, para estar pronto daqui a cerca de dois anos. Esta iniciativa tem como objectivo promover e recompensar actividades de investigação e desenvolvimento que tenham impacto nos sistemas de transporte terrestre [33].

Esta competição irá decorrer durante 30 meses de forma faseada:

- Fase 1 (6 meses): Design e conceitos iniciais
- Fase 2 (1 ano): Competição virtual
- Fase 3 (1 ano): Implementação e prototipagem

No fim da primeira fase apenas 4 equipas serão seleccionadas, recebendo um prémio de 6000 € (distribuído pelas 4). Na segunda fase a equipa vencedora receberá 5000 € e na última etapa a equipa vencedora receberá 40000 € de prémio.

Existem ainda ajudas para construção dos protótipos de cada equipa: 18000 € e 21000 € a distribuir pelas 4 equipas, na segunda e terceira fases do concurso, respectivamente.

O protótipo a desenvolver tem que assegurar o transporte de 9 pessoas num ambiente urbano durante cerca de 2 quilómetros, respeitando as regras de trânsito e de segurança, de forma autónoma. Para completar com sucesso este desafio, o veículo terá de ser capaz de passar com sucesso em duas provas, uma de dia e outra de noite, partindo de uma estação inicial, recolhendo e distribuindo os seus passageiros por outros sítios de paragem pré-definidos. Estas provas têm que ser feitas tão rápido quanto possível, dentro dos limites, e usando a menor energia possível.

Para cumprir as condições de participação é necessário que as equipas tenham no máximo 10 elementos e que pelo menos um destes elementos esteja a realizar de momento uma tese de doutoramento.

Para facilitar o trabalho das equipas é possível a utilização de uma plataforma já desenvolvida pela Critical Move, sendo também possível a construção de raiz de uma nova plataforma pelas equipas.

2.4.3 CyberCar

O excesso de tráfego nas cidades europeias tem vindo a aumentar de forma muito acentuada nas últimas décadas. Como forma de resolver este problema, enormes avanços tecnológicos têm sido registados no sector dos transportes.

O projecto europeu CyberMove [34] procurou abordar esta problemática desenvolvendo e promovendo sistemas de transporte alternativos baseados em veículos eléctricos autónomos (sem condutor), denominados de CyberCars.

A Universidade de Coimbra no âmbito deste mesmo projecto, desenvolveu através do Laboratório de Robótica Móvel do Instituto de Sistemas e Robótica [35] um veículo eléctrico para actuar em ambientes urbanos, um CyberCar. Este veículo foi desenvolvido numa plataforma da Yamanha, parceira deste projecto (ver figura 2.17).

Este veículo foi desenvolvido para oferecer um meio de transporte mais flexível, seguro e económico, usando energia limpa, por forma a completar os sistemas de transporte principais de uma forma mais personalizada. Desta forma é possível por exemplo passear por uma zona histórica a bordo destes veículos, sem se preocupar com congestionamento ou restrições de circulação aos veículos comuns.

Existe uma forte componente científica neste projecto, que incide em três áreas chave: navegação autónoma, sistemas de anti-colisão e interfaces avançados com o utilizador.

O CyberCar é um veículo controlado por computador e assistido por tecnologias de comunicação sem fios, tais como o GPS. Embora circule sem intervenção humana, este sistema é capaz de respeitar



Figura 2.17: Plataforma de um CyberCar da Yamaha.

os sinais de trânsito, parando quando se encontra na presença de obstáculos ou pessoas, respeitando os sítios de paragem dos passageiros e voltando ao ponto de origem quando não é mais necessário. Para fazer a detecção de obstáculos, este veículo é munido de sensores de ultra-sons e lasers, como forma de detecção de qualquer objecto que intercepte a sua trajectória.

A sua trajectória é definida por um fio eléctrico colocado debaixo do solo ou à superfície, para que possa ser possível seguir o seu campo magnético, servindo assim de guia aos veículos. Este tipo de técnica para condução de veículos é definida como guidewire ou veículos filoguiados, ou seja, seguem um percurso pré-definido por um fio eléctrico.

Quando é necessária a sua utilização por um passageiro, basta aguardar pela sua passagem ou chamá-lo através de um PDA. Desta forma é possível obter um nível mais personalizado de transporte, diminuindo as emissões de gases, poluição sonora e congestionamento do tráfego.

2.4.4 Busca e salvamento

Um dos objectivos da criação de elementos robóticos autónomos é aceder a locais perigosos ou de difícil acesso aos humanos. O Laboratório de Sistemas Autónomos do Instituto Superior de Engenharia do Porto desenvolveu alguns projectos relacionados com monitorização ambiental e exploração remota, entre eles o ROAZ e o FALCOS (ver figura 2.18) [36].



Figura 2.18: Da esquerda para direita, os robôs dos projectos ROAZ e FALCOS [36].

O ROAZ é um projecto criado para ambientes aquáticos, rios e oceano, para operações de oceanografia, batimetria³ e busca e salvamento. O veículo está equipado com uma antena WIFI, uma ligação sem fios de 2.4Ghz para transmissão de vídeo, GPS e equipamento de visão. O equipamento de visão instalado é composto por uma câmara com capacidades de visão nocturna e uma câmara termográfica de infravermelhos [36].

³ciência que se ocupa do estudo da topografia e profundidade dos oceanos

O FALCOS é um sistema de pequena dimensão projectado para operar a baixa altitude em missões de prevenção de fogos florestais, captando imagens aéreas. É um avião robótico não tripulado, equipado com tecnologia capaz de detectar pequenos fogos a 300 metros de altitude. Este avião possui uma câmara convencional e uma outra que capta frequências próximas do infravermelho, o que permite captar a chama de um isqueiro a 3 metros de distância [36].

Capítulo 3

Visão computadorizada

Os órgãos sensoriais representam num veículo a sua forma de perceber o mundo, constituindo uma das partes mais importantes na construção de um veículo autónomo. Existe um vasto número de equipamentos sensoriais possíveis de usar num sistema deste tipo, dependendo das necessidades e finalidade a que o veículo se destina e do orçamento disponível.

Existem equipamentos de medição directa que adquirem um tipo de informação específica, como é o exemplo de sensores de distância ou movimento, e sistemas de percepção que permitem adquirir um tipo de informação mais abstracta, tal como é o caso de sistemas de aquisição de imagem.

Os sensores utilizados nos veículos para condução autónoma são normalmente divididos em duas categorias: activos e passivos. Os sensores laser, ultra-sons, infra-vermelhos, GPS e outros mais, são referidos como activos uma vez que medem o tempo de “viagem” de um sinal entre a sua emissão e recepção, após ser reflectido por qualquer objecto. A maior vantagem deste tipo de sensores é que conseguem efectuar medidas, normalmente distâncias, de uma forma directa e computacionalmente pouco pesada. No entanto, este tipo de sensores têm também desvantagens tais como o facto de não oferecerem por vezes uma grande resolução espacial, baixa velocidade em certos casos e as interferências entre sensores que podem acontecer caso estes se depararem com uma situação confusa em que circulem vários veículos [37].

Os sensores ópticos de aquisição de imagem são referidos como passivos, uma vez que são capazes de adquirir informação de uma forma “não intrusiva”. Com a banalização deste tipo de equipamentos é possível actualmente incorporar vários destes dispositivos num veículo, por forma a possibilitar um campo de visão bastante abrangente [38].

A visão computadorizada pode ser entendida como a capacidade de os computadores perceberem o conteúdo de uma imagem, ou sequências de imagens. Esta percepção significa a extracção de certas características pretendidas numa imagem, para um determinado fim, de acordo com a tarefa a desempenhar. Por exemplo, a visão computadorizada pode ser utilizada numa linha de montagem, ou em veículos autónomos, analisando as imagens para que possam ser posteriormente tomadas as respectivas acções, ou então pode ser usada apenas para efeitos de visualização, em análise de superfícies ou materiais para detectar certas características que poderiam não ser detectadas a olho humano.

Na área da condução autónoma a visão computadorizada tornou-se nos últimos anos o método de percepção mais estudado, uma vez que é o que revela mais capacidades. A informação sensorial proveniente de uma câmara é talvez a mais rica, mas no entanto é também a mais abstracta e de mais difícil compreensão. Existem na imagem muitas condicionantes que podem dificultar a tarefa de compreender o seu conteúdo tais como a luminosidade, variabilidade de formas e cores, complexidade de ambientes. Estas condicionantes resultam muitas vezes numa dificuldade acrescida no processo de

separação entre informação útil e ruído.

A aplicação de métodos de análise visual em veículos para efeitos de condução autónoma depende efectivamente do fim a que se destina o veículo e em que ambientes irá funcionar. Um veículo que é desenvolvido para condução autónoma em auto-estrada irá certamente ter bastantes dificuldades em navegar em ambientes que não sejam tão estruturados, tanto no piso como nos limites de estrada. Um veículo desenhado para navegar em todo-o-terreno irá sentir iguais dificuldades em auto-estrada. Um exemplo destes dois tipos de ambiente são as imagens da figura 3.1 que representam exemplos de um ambiente estruturado e não estruturado.

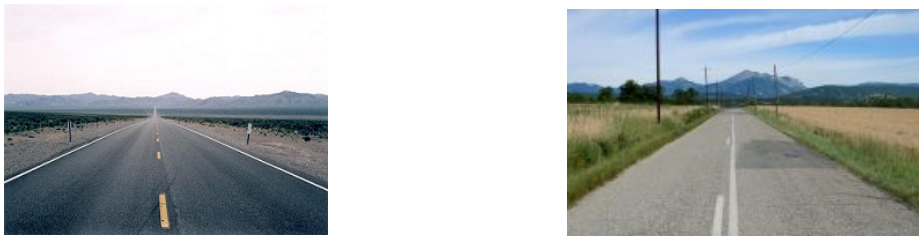


Figura 3.1: Exemplos de ambientes estruturados (à esquerda) e não estruturados (à direita).

Pode-se então falar de duas grandes categorias de ambientes: os estruturados e os não estruturados. Em ambientes estruturados, os sistemas de visão dependem da própria estrutura da estrada, orientando-se assim por características tais como as marcas limitadoras, cor ou a própria simetria da faixa de rodagem. Os sistemas desenvolvidos para navegação em ambientes não estruturados não podem ter como suporte este tipo de características de uma faixa de rodagem estruturada, uma vez que estas podem não existir. Existe neste caso uma tarefa muito mais complexa por parte do sistema para tentar perceber em qualquer altura qual poderá ser a rota a seguir [39].

3.1 Aquisição de imagem

Existem vários tipos de sistemas de percepção visual possíveis de utilizar, com diferentes capacidades, de acordo com as necessidades do próprio sistema que irá processar a informação adquirida. Existem sistemas que necessitam de obter informação visual em todas as direcções, tal como existem outros que apenas necessitam de capturar imagens em apenas uma direcção. Existem também sistemas que necessitam de perceber o ambiente em três dimensões e outros ainda que apenas necessitam de duas dimensões.

Podem-se distribuir os sistemas de aquisição de imagem em categorias: omni-direccionais, estereoscópicos ou monoculares.

Os sistemas omni-direccionais, tal como o nome indica, são capazes de capturar imagem em todas as direcções, no entanto, a deformação introduzida na imagem resultante do efeito perspectiva torna estes sistemas não muito comuns em condução autónoma. Na imagem do lado esquerdo da figura 3.2 é possível ver um exemplo de um setup de aquisição de imagem omni-directional, estando do lado direito representada uma imagem adquirida com um sistema deste tipo.

Existem também sistemas monoculares, constituídos apenas por uma câmara, que permitem observar o mundo em duas dimensões. Este tipo de sistema é utilizado para ambientes menos complexos nos quais não é necessário ter uma noção das três dimensões, no entanto já é possível fazer uma análise relativamente completa do ambiente. Na figura 3.3 podemos ver o setup comum de um sistema de aquisição monocular.



Figura 3.2: Sistema de visão omni-direcional e exemplo de aquisição de imagem segundo este mesmo sistema [40, 41].

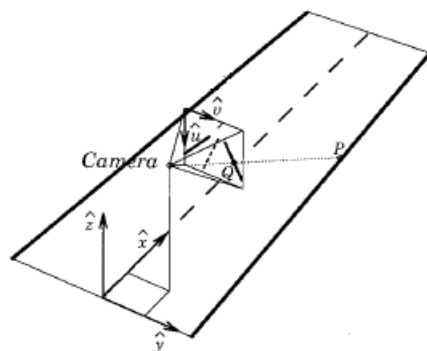


Figura 3.3: Sistema de visão monocular [40].

Bastante utilizados são também os sistemas estereoscópicos, compostos por duas câmaras, que permitem desta forma ter uma noção de três dimensões. Estes sistemas são utilizados quando as necessidades de aquisição da informação são mais complexas. Na figura 3.4 é possível ver a representação de um sistema de aquisição estereoscópico bem como o exemplo da aquisição de duas imagens, pelas câmaras esquerda e direita.

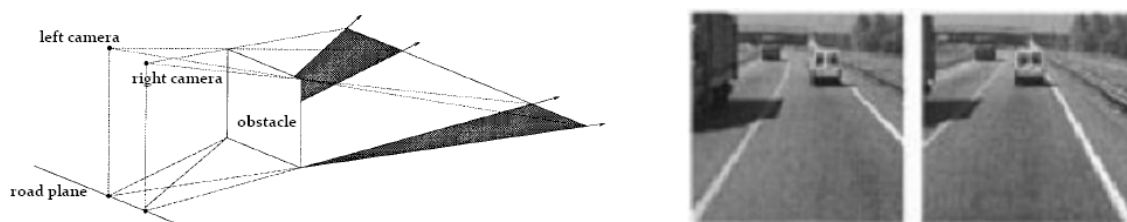


Figura 3.4: Sistema de visão estereoscópico e exemplo de aquisição de imagem segundo este mesmo sistema [40].

3.2 Espaços de cor

A cor é a forma como o sistema visual humano consegue medir ou perceber parte do espectro electromagnético. Devido às propriedades deste mesmo sistema, não somos capazes de ver todas as combinações possíveis do espectro visível, mas tendemos a agrupar vários espectros em cores.

Um espaço de cor é uma notação para representar uma determinada cor (crominância) e uma determinada intensidade luminosa (luminância). Existem diversos espaços de cor, sendo alguns deles

mais eficientes no ponto de vista de percepção (aproximam-se do sistema visual humano) e outros mais eficientes no ponto de vista tecnológico (são mais fáceis de representar e transmitir).

Os espaços de cor podem por isso ser divididos em várias categorias de acordo com a sua finalidade ou utilização. Existem vários espaços de cor que foram criados devido à necessidade específica de certos sistemas, tais como a TV (YUV, RGB), ou estão relacionados com dispositivos tais como as impressoras (CMY) ou câmaras fotográficas. Existem também espaços de cor que estão relacionados com o sistema visual humano, tentando aproximar-se das suas propriedades para aumentar a percepção e compreensão destes sistemas (HSV) [42, 43].

Os espaços de cor RGB e HSV receberão mais atenção neste trabalho, sendo feita de seguida uma explicação mais detalhada sobre eles, uma vez que os outros espaços de cor não são considerados relevantes para este trabalho.

3.2.1 RGB

O espaço de cor RGB (Red, Green, Blue) é um sistema aditivo no qual o vermelho, o verde e o azul são combinados para reproduzir outras cores. Este modelo define-se como aditivo uma vez que resulta da sobreposição de projecções de luz das cores primárias com diferentes intensidades (ver figura 3.5).

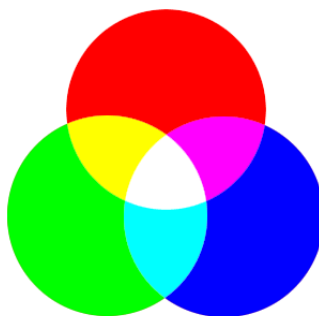


Figura 3.5: Espaço de cores RGB.

Neste espaço de cor os píxeis de uma determinada imagem são representados através de três números que indicam a proporção de cada componente de vermelho, verde e azul. Cada uma pode variar entre o mínimo (completamente escuro) e máximo (completamente intenso). Quando todas as componentes estão no mínimo, o resultado é preto. Se todas estão no máximo, o resultado é branco. Qualquer cor pode ser criada através da combinação do vermelho, verde e azul.

Na figura 3.6 podemos visualizar uma imagem representada em RGB e a correspondente divisão nas suas três componentes. Como exemplo, é possível verificar que o papagaio da direita é em grande parte vermelho e como tal se analisarmos a divisão em componentes RGB em baixo, verificamos que a componente vermelha nessa zona é mais clara do que nas outras componentes. Ou seja, esta zona é maioritariamente constituída pela primeira componente, o vermelho.

3.2.2 HSV

Este modelo é normalmente utilizado em aplicações que envolvem a definição de cor por parte do homem, uma vez que a sua estrutura se baseia na forma como o olho humano percebe a cor. É mais fácil e intuitivo, por parte de um humano, conseguir atingir uma determinada cor modelando os parâmetros do modelo HSV do que utilizando o modelo RGB.

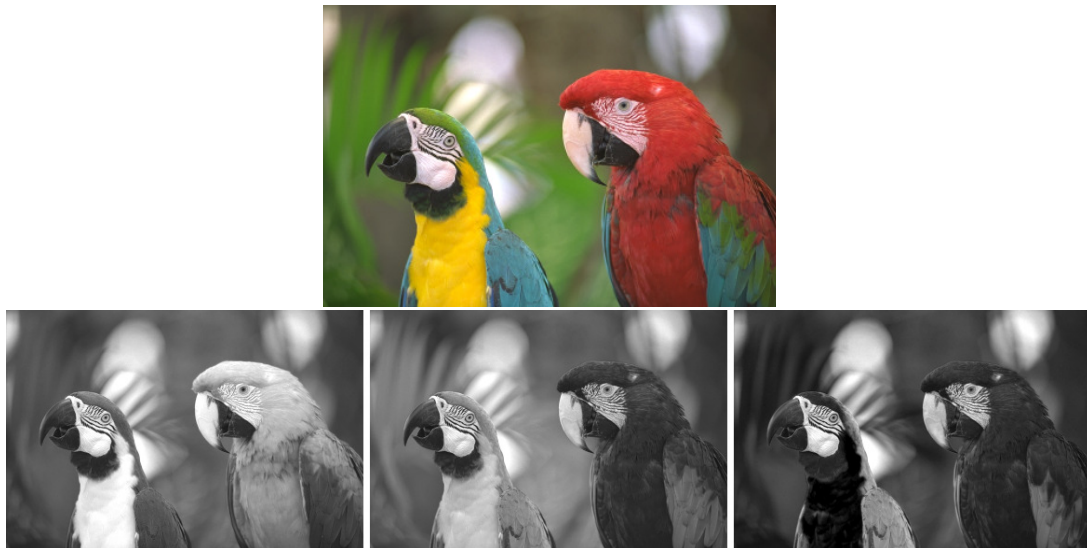


Figura 3.6: Divisão de uma imagem em RGB (em cima) nas suas 3 componentes (em baixo) [44].

O espaço de cor HSV é composto por três componentes: Hue (matiz/tonalidade), Saturation (saturação) e Value (intensidade/brilho). O seu espaço de cor é usualmente representado por um cone (ver figura 3.7).

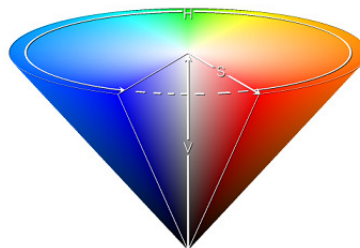


Figura 3.7: Espaço de cor HSV.

A matiz pode ser definida como a característica que descreve a cor, abrangendo todas as cores do espectro visível. Mais especificamente representa em termos espectrais o comprimento de onda dominante. É esta componente que define se a cor é vermelha, verde, amarela, etc.

A saturação refere-se ao nível de pureza da cor, ou à “quantidade de cor”. Quanto menor for esta componente mais a cor se aproxima do cinzento. Ao nível espectral a saturação define-se pela dispersão da cor em redor do comprimento de onda dominante. Quanto menos dispersão existir em torno deste comprimento de onda dominante, ou seja do valor da matiz, mais saturada será esta cor.

O brilho tal como o nome indica refere-se ao brilho desta mesma cor. Representa a quantidade de luz que é percebida e posteriormente reflectida por um determinado objecto. Um valor elevado de brilho aproxima a cor do branco e um valor baixo aproxima a cor do preto [31, 42].

Normalmente quando é necessário especificar uma determinada cor no modelo HSV, o mais usual será aparecer algo semelhante ao representado na figura 3.8. Desta forma a matiz é representada por uma região circular representando todas as tonalidades possíveis, sendo depois possível regular a saturação e o brilho no triângulo que se encontra no interior.

Este modelo revelou-se bastante interessante no trabalho realizado, uma vez que após uma expe-

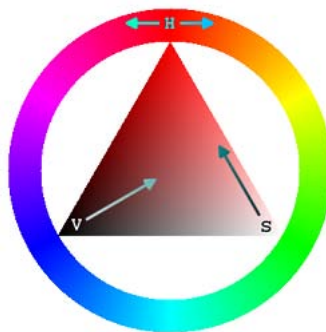


Figura 3.8: Espaço de cor HSV.

riência inicial com o modelo RGB, o HSV revelou-se muito mais intuitivo na forma de perceber a constituição das imagens analisadas.

3.3 Pré-processamento da imagem

O processo de aquisição e interpretação de informação visual requer por vezes certos procedimentos que têm como objectivo melhorar ou realçar as características que se pretendem detectar. Outros procedimentos são também utilizados em situações em que existe um conjunto de informação demasiado grande para analisar, tornando-se imperativo simplificar este conjunto de informação para simplificar o processo de análise.

Este conjunto de procedimentos, que antecede a análise propriamente dita, podem ser definidos como a parte de pré-processamento da imagem. Assim, de forma geral as técnicas de pré-processamento têm como principal objectivo a simplificação das imagens, realçando as suas características mais importantes.

Neste caso, quando falamos de condução autónoma, podemos referir processos que após a aquisição da imagem tentam realçar as características da estrada, dos seus limites, ou até mesmo diferenciar obstáculos que nela se encontrem.

Nesta secção serão apresentadas três técnicas comuns de pré-processamento de imagem: extracção de contornos, segmentação de cor e transformações morfológicas.

3.3.1 Extracção de contornos

A detecção de contornos torna-se uma tarefa essencial na generalidade dos sistemas de visão computorizada, uma vez que são os contornos que definem os limites entre os vários objectos presentes na imagem, sendo por si só um bom método para extrair os componentes estruturais mais importantes da imagem.

Contornos são os limites que separam regiões com diferentes brilhos ou cores. Definem-se na prática pelas zonas na imagem onde o valor da intensidade muda bruscamente, ou seja, onde o valor do gradiente entre pixéis que se encontram juntos é elevado (em imagens monocromáticas).

Os contornos definem assim normalmente os limites de um objecto, ou as marcas mais visíveis que pertencem a este mesmo objecto. No entanto, uma vez que uma imagem é uma representação bi-dimensional de uma cena tri-dimensional, um contorno pode também representar uma descontinuidade na profundidade da cena tri-dimensional representada pela imagem.



Figura 3.9: Processo de extracção de contornos [45].

A figura 3.9 representa a aplicação de um algoritmo de extracção de contornos a uma imagem. Tal como é possível verificar, através da aplicação deste algoritmo são extraídas todas as características estruturais mais importantes da imagem, simplificando assim a sua análise posterior.

Um detector de contornos deve ser capaz de marcar numa imagem tantos contornos quanto possível, na sua localização correcta e com tanta imunidade ao ruído quanto possível.

O algoritmo de Canny é talvez o algoritmo de extracção de contornos mais popular. Este algoritmo actua sobre imagens monocromáticas, devolvendo como resultado imagens binárias onde os píxeis com valores diferentes de zero marcam a localização de contornos. Os parâmetros deste algoritmo permitem especificar o tipo de processamento desejável, se se quer encontrar todos os contornos ou apenas os mais definidos, definindo dois níveis de threshold, um alto e outro baixo.

O algoritmo de Canny processa a imagem por forma a encontrar todos os contornos possíveis, em todas as direcções. No entanto, por vezes é necessário fazer uma detecção de contornos que tenham uma determinada orientação. O algoritmo de Sobel é uma técnica de detecção de contornos horizontais e verticais, que permite desta forma encontrar apenas os contornos que nos interessam.

Apenas é possível aplicar o algoritmo de Sobel a imagens com apenas um canal, monocromáticas, sendo que a imagem resultante é binária, onde, à semelhança do algoritmo de Canny, os píxeis com valores diferentes de zero marcam a localização dos contornos.

3.3.2 Segmentação de cor

A tarefa de segmentação de cor numa imagem define-se pelo processo de isolar zonas ou regiões que contenham a mesma cor, ou características de cor muito semelhantes. O objectivo final é criar uma divisão de regiões homogéneas que se caracterizem por terem características de cor e textura muito semelhantes, por forma a identificar objectos ou regiões específicas.

Normalmente este processo é utilizado para identificar regiões específicas, com características pré-conhecidas, por forma a isolar estas regiões eliminando as restantes zonas da imagem que não tenham estas mesmas características. Em sistemas de condução autónoma normalmente este processo é utilizado por forma a isolar a faixa de rodagem ou objectos com uma cor específica.

Em casos em que o ambiente é não estruturado, a segmentação da imagem pode ser feita com o objectivo de classificar as zonas da imagem como estrada ou não estrada. Na figura 3.10 é possível verificar que as imagens de baixo resultam da segmentação das imagens de cima. Esta segmentação é resultado da análise das características da estrada, eliminando tudo o resto na imagem que não possua as mesmas características.

Nos casos em que se trata de ambientes estruturados, a segmentação da imagem poderá ser por exemplo usada para realçar as faixas brancas que delimitam a faixa de rodagem. Neste caso o pré-



Figura 3.10: Exemplo de segmentação [46].

processamento da imagem pode por exemplo ser baseado no conhecimento prévio das texturas ou cor esperada, geometria provável da estrada ou baseado na extracção de faixas brancas caso estas existam.

3.3.3 Operações morfológicas

As operações morfológicas em imagem definem-se pela análise e processamento de estruturas geométricas, sendo normalmente utilizadas para remoção de ruído e isolamento ou junção de regiões com determinadas características. De forma geral, tendo em atenção o trabalho a ser desenvolvido, as operações morfológicas de interesse são a dilatação e a erosão de grupos de píxeis, em imagens monocromáticas.

A dilatação, tal como o nome indica, é usada para expandir as zonas constituídas por píxeis com alto valor de brilho, resultando no aumento das regiões mais claras da imagem fazendo com que se espalhem para outras regiões. Assim é possível por exemplo numa imagem de uma estrada conseguir aumentar o realce das faixas brancas delimitadoras, caso estas existam (ver imagem do meio da figura 3.11).

A erosão é o processo inverso, ou seja, tem como função encolher os píxeis com valor de brilho elevado, isolando as regiões mais claras da imagem. Pode ser usado por exemplo para eliminar o ruído presente numa imagem (ver imagem da direita da figura 3.11).

A combinação destes dois processos pode também ser utilizada para aumentar o realce de certas formas, que devido a determinadas circunstâncias como excesso de luminosidade, não estejam correctamente representadas na imagem adquirida.

3.4 Análise da imagem

Uma das primeiras tarefas de um qualquer veículo que pretende ser autónomo é desenvolver capacidades de detecção dos limites da faixa de rodagem, bem como o seu posterior seguimento. Várias



Figura 3.11: Da esquerda para a direita: a imagem original, aplicação do processo de dilatação e aplicação do processo de erosão.

abordagens podem ser seguidas, tais como baseadas no seguimento das linhas limitadores das faixas, em casos que o ambiente é estruturado, ou baseadas em análise de texturas ou regiões [17].

Os sistemas que seguem a abordagem baseada nas linhas limitadoras das faixas, assumem que todas as estradas têm estas propriedades, o que na realidade não é verdade, ou seja, podem ser bastante falíveis.

Os sistemas de navegação autónoma que se baseiam na análise de regiões e texturas para perceber a localização e limites da estrada conseguem ser mais flexíveis, quanto ao tipo de estrada em que conseguem conduzir. Esta abordagem analisa a imagem tendo como objectivo classificar as diferentes regiões como estrada ou não estrada. Estes sistemas não necessitam de estradas completamente estruturadas para funcionar, mas no entanto são muito mais dependentes das condições de iluminação. Uma vez que a cor dominante será o cinzento, as sombras ou luz do sol podem causar bastantes problemas no funcionamento deste tipo de algoritmos.

A solução ideal para conseguir uma boa detecção da zona de estrada, em qualquer ambiente quer seja estruturado ou não, deve ser indiferente a condições de iluminação ou visibilidade, o que é portanto muito difícil de alcançar. Este objectivo é só possível de alcançar com uma integração de informações pré-conhecidas, tais a informação geométrica e cromática esperada numa estrada, com informação dinamicamente adquirida pelo veículo em cada momento, para que o sistema se possa adaptar a novas situações. Trata-se de aproveitar todos os factores possíveis para se convergir numa solução com um nível de confiança aceitável. Por exemplo, com base na geometria esperada da estrada é possível condicionar a forma como é feita a percepção da textura da estrada, tornando assim o processo de segmentação mais eficaz.

No caso da detecção e possível seguimento de obstáculos, a lógica é um pouco semelhante. As abordagens podem também ser várias, mas o procedimento usual para tentar detectar obstáculos é, após ter definido uma zona de estrada, tentar verificar se algum objecto se encontra nessa mesma zona.

Enquanto a estrada é tipicamente vista num plano, os obstáculos têm altura. A utilização de duas câmaras potencia neste caso a identificação dos obstáculos a três dimensões, tirando proveito das diferenças entre as duas imagens capturadas.

Existem diversas metodologias para a detecção de obstáculos, mas de uma forma simplificada, de acordo com o conceito referido em [47], a detecção de obstáculos segue o seguinte procedimento:

1. Detecção de possíveis candidatos a obstáculos
2. Classificação dos candidatos em falsos alarmes ou obstáculo reais
3. Seguimento dos obstáculos reais

3.5 Técnicas de análise de imagem

Nesta secção faz-se uma abordagem genérica a várias metodologias ou técnicas de análise de imagem, que foram consideradas importantes no contexto deste trabalho. As técnicas aqui descritas foram escolhidas por forma a abordar vários métodos com características muito distintas para dar uma ideia geral das várias possibilidades.

Começar-se-á pelo optical flow que é uma técnica que através da percepção da dinâmica do movimento do veículo, tenta extrair alterações a esta mesma dinâmica para gerar hipóteses da localização de obstáculos.

Serão também descritas, de forma geral, as técnicas que usam modelos pré-definidos para depois tentar comparar estes modelos com alguns dos objectos que encontrem na imagem, para efectuar a detecção de veículos.

De seguida irá ser descrito o Ralph, que é um sistema concebido para suporte à condução autónoma em ambientes estruturados, usando uma região de análise auto-adaptável que se ajusta de acordo com a dinâmica do veículo.

Ao contrário do RALPH, existem também sistemas que não necessitam de ambientes estruturados para efectuar condução autónoma, como é o caso do UNSCARF. O UNSCARF é um sistema que tem como objectivo a condução autónoma em ambientes não estruturados, utilizando uma “nova” metodologia de segmentação da estrada.

Para dar um exemplo de um sistema que utilize um sistema de aquisição estereoscópico, irá ser referido o GOLD. Este sistema utiliza duas câmaras para efectuar condução autónoma com detecção de obstáculos.

3.5.1 Optical flow

Nesta secção irá ser feita uma abordagem genérica ao optical flow, com objectivo de se entender as características deste método bem como as suas capacidades e limitações quando aplicado a certas situações. Uma abordagem mais detalhada deste método pode ser encontrada em [48, 49].

Opostamente aos sistemas que utilizam visão estereoscópica, os sistemas que utilizam apenas uma câmara necessitam de informação adicional para tentar reconstruir a sua percepção de longitudinalidade, uma vez que nas imagens adquiridas apenas é possível ter percepção de duas dimensões. Este método é geralmente usado apenas para detecção de obstáculos.

O método de optical flow pode ser descrito como a avaliação da diferença entre imagens consecutivas, numa sequência de imagens em escala de cinzentos, que pode de certa forma ser utilizada para perceber os aspectos dinâmicos desta mesma sequência. A utilização deste tipo de algoritmos permite retirar informações não só da diferença temporal entre imagens, como também da própria estrutura espacial da sequência de imagens analisada [50].

O procedimento básico para proceder à detecção de obstáculos que se encontrem à frente de um veículo, usando optical flow, resume-se aos seguintes passos:

1. Cálculo dos vectores de optical flow. Estes vectores ligam uma região projectada numa imagem com a posição desta mesma região na imagem imediatamente seguinte.
2. Cálculo de uma estimativa destes mesmos vectores de optical flow, assumindo que não existe nenhum obstáculo e que a translação da câmara é feita de forma planar. Estes vectores tentam estimar o deslocamento de pontos entre imagens consecutivas, que correspondem ao deslocamento dos pontos da cena, projectados nestas mesmas imagens.

3. Avaliar a diferença entre os vectores calculados e os vectores estimados.



Figura 3.12: Cálculo dos vectores de optical flow: (a) duas imagens consecutivas; (b) os vectores de optical flow.

O cálculo dos vectores de optical flow pode ser feito de diversas formas, recorrendo a várias técnicas, mas sempre com o objectivo de ligar uma determinada região de uma imagem com essa mesma região da imagem seguinte. Normalmente este processo é feito utilizando imagens monocromáticas, tentando fazer uma detecção tão eficiente quanto possível, no menor tempo possível. Para melhor descrição dos métodos possíveis para o cálculo dos vectores de optical flow ler [49].

Os vectores estimados são determinados através de modelos pré-definidos que têm em conta a dinâmica do veículo e os parâmetros da câmara, tentando prever a deslocação de cada ponto de uma cena projectado numa imagem. Num ambiente perfeito em que fosse possível obter um modelo exacto da dinâmica do veículo, estes vectores estimados seriam exactamente iguais aos calculados através da análise das imagens, em casos em que não existam veículos a circular na estrada.

É sobre este modelo que se baseiam os métodos de optical flow, mas os algoritmos que utilizam este método podem conter muitas variações [51]. No 1º passo, cálculo dos vectores, é possível introduzir várias condicionantes, tais como restringir o tamanho máximo dos vectores ou limitar a região da imagem na qual os vectores são calculados. Existem também alternativas na escolha dos pontos que irão originar os vectores. Pode-se optar por localizar os vectores nos pontos que pertencem aos contornos de uma imagem, variando assim a sua posição em cada imagem ou pode-se ainda dividir a imagem num reticulado e colocar os vectores no centro de cada quadrado (ver figura 3.12).

Para detecção de obstáculos, todas as diferenças entre vectores calculados e estimados têm que ser calculadas. Um possível obstáculo é detectado se se verificar uma zona em que a diferença dos vectores seja realmente notória. Se não existirem quaisquer obstáculos, as diferenças entre vectores são aproximadamente nulas em toda a imagem [8].

O optical flow apresenta no entanto contrapartidas:

1. Elevado peso computacional.
2. Baixa precisão.
3. Necessidade de obter um bom modelo dinâmico para fazer uma estimativa precisa dos vectores.

O elevado peso computacional deve-se ao cálculo dos vectores de optical flow. O cálculo destes vectores, seja qual for a técnica utilizada, necessita de fazer várias comparações para cada vector, por forma a verificar qual será a região na frame seguinte que mais se assemelha a uma região na frame actual.

A baixa precisão deste método deve-se ao facto de ser difícil compreender por absoluto estruturas tri-dimensionais e porque os vectores calculados podem não corresponder exactamente à deslocação de uma determinada região. Nos casos de regiões que se encontrem longe do dispositivo de aquisição é também difícil fazer correspondência entre imagens uma vez que quanto mais longe estiver uma região, menos píxeis na imagem a representam, devido ao efeito perspectiva.

A utilização apenas de optical flow para efectuar análise de imagens para detecção de obstáculos não é muito utilizada. É no entanto uma metodologia que utilizada em conjunto com outros órgãos sensoriais, ou outras metodologias de análise de imagem, pode resultar em sistemas bastante interessantes.

A sua utilização em sistemas de vigilância ou controlo de tráfego é talvez mais proveitosa. Neste tipo de situações não é necessário um nível tão elevado de precisão, bem como o facto de o equipamento de aquisição estar estático, o que possibilita assim melhores resultados, conseguindo detectar com mais precisão certas mudanças no ambiente que se a destina analisar.

3.5.2 Detecção baseada em modelos

Nesta secção irá ser feita uma descrição dos algoritmos de detecção de obstáculos que se baseiam na utilização de modelos pré-definidos (model based detection), descrevendo-se apenas a abordagem global e as metodologias utilizadas.

Esta abordagem define-se pela criação de um modelo, que contenha tanta informação estrutural quanto possível do objecto a detectar. Desta forma são exploradas as propriedades bem conhecidas dos objectos, usando esta informação por forma a encontrar nas imagens itens que contenham tanto quanto possível estas mesmas propriedades. Neste caso os objectos a detectar são veículos, e é sobre este caso específico que irá ser descrita a abordagem comum a este problema. Normalmente são definidos modelos geométricos dos veículos tentando representar estas mesmas propriedades comuns a todos dos veículos, para que seja possível corresponder tantos veículos quanto possível a estes mesmos modelos [52, 53]. As imagens da figura 3.13 representam a definição de um modelo geométrico que se espera ser semelhante à generalidade dos veículos, para que a correspondência do modelo com os veículos numa situação real possa ter bons resultados.

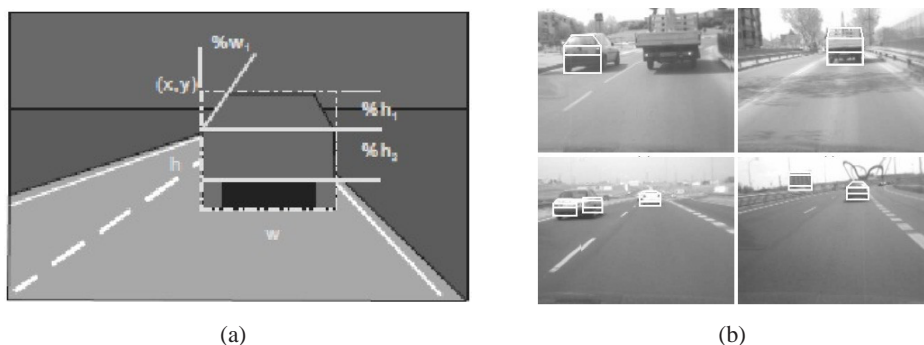


Figura 3.13: Utilização de um modelo geométrico para detecção de veículos: (a) definição do modelo; (b) correspondência do modelo com veículos reais [52].

Em geral, um algoritmo de detecção baseado em modelos necessita de:

1. O modelo inicial

2. O modelo de deformação, que permite calcular as alterações no modelo inicial de acordo com a sua posição
3. Um algoritmo de pesquisa para encontrar a melhor correspondência do modelo em qualquer item da imagem

A figura 3.14 representa um mecanismo genérico de detecção de veículos utilizando a correspondência com modelos. Neste caso é feita uma extracção inicial das características da imagem que normalmente se encontram presentes num veículo. De seguida é feita uma comparação da ocorrência destas características com os modelos de veículos que foram previamente gerados e guardados numa base de dados, por forma a tentar encontrar correspondências que indiquem presença de veículos.

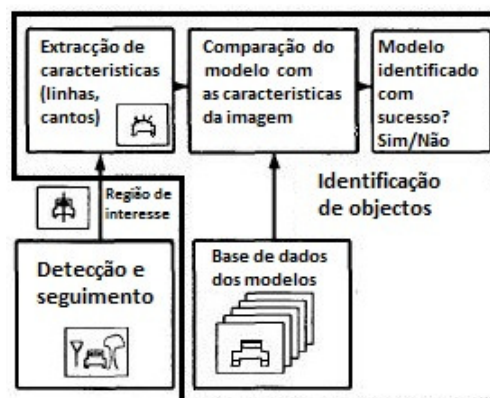


Figura 3.14: Detecção de veículos baseada em modelos.

Existem várias características dos veículos que podem ser aproveitadas para a construção de modelos eficazes para a sua detecção: simetria, cor, sombras, cantos, contornos verticais/horizontais, texturas, as luzes dos veículos, etc [54].

A simetria é uma característica presente na generalidade dos veículos, sendo frequentemente usada para detecção e reconhecimento de veículos em aplicações que utilizam visão computadorizada. Imagens de veículos observados pela frente ou pela retaguarda apresentam normalmente simetria vertical e horizontal. Várias abordagens fazem uso desta simetria para calcular os candidatos a veículos [54, 55].

Os cantos [56] e os contornos verticais/horizontais são características presentes em todos os veículos que pode ser normalmente aproveitadas, tal como mostrado pelas figuras 3.15 e 3.16. Este tipo de abordagem explora o facto de os veículos normalmente serem constituídos por 4 cantos, formando uma caixa rectangular. Além disso um veículo apresenta sempre várias estruturas que contêm contornos verticais e horizontais, tais como os vidros, pára-choques, matrícula, etc. A detecção deste tipo de características torna-se assim uma forte possibilidade a explorar neste tipo de sistemas de detecção de veículos baseados em modelos.

Embora existam poucas abordagens que utilizem a cor e a textura para a detecção de obstáculos, estas propriedades podem ser usadas como informação complementar para esta mesma tarefa. Estas propriedades são de difícil utilização uma vez que não existe uma cor ou textura comum a todos os veículos.

Existem no entanto algumas abordagens que tentam segmentar a cor da estrada e desta forma tentar encontrar possíveis objectos que se encontrem nesta zona de estrada, que tenham cor diferente da cor da própria estrada. A figura 3.17 é um exemplo de um destes tipos de abordagem. Neste caso

é utilizada a intensidade de cada píxel. É traçada uma linha na imagem que tem como função criar um perfil da intensidade da imagem naquela zona, com o objectivo de encontrar zonas com valores semelhantes.

Devido a sombras, condições climáticas ou outras condições que afectem a visibilidade, é necessário construir um modelo com tanta informação quanto possível. No entanto é também necessário que o modelo seja tão genérico quanto possível, para que possa ser adaptável a qualquer veículo.

Os algoritmos de detecção de obstáculos baseados em modelos apresentam resultados satisfatórios em ambientes relativamente estruturados onde não exista muita sobreposição de obstáculos na imagem. Um dos factores que influencia o sucesso destes métodos é a capacidade de após ter detectado um veículo, conseguir seguir o seu movimento com relativa facilidade. Um veículo detectado numa frame poderá mudar de posição e tamanho na frame seguinte, devido ao seu movimento, sendo por isso necessário compreender a sua deformação de acordo com a distância a que este se encontra do ponto de aquisição de imagem. Para isso é necessário construir um modelo dinâmico que seja capaz de prever com relativa fiabilidade a localização de um veículo na próxima frame, após este ter sido detectado. Várias técnicas foram já desenvolvidas para lidar com a deformação dinâmica dos modelos, de forma a tentar resolver este problema, tal como a referida em [57].

3.5.3 RALPH: Rapidly Adapting Lateral Position Handler

O RALPH, desenvolvido ao abrigo do projecto CMU-Navlab (ver secção 2.2.2), para apoio à condução autónoma em ambientes estruturados, utilizou uma abordagem interessante e inovadora na análise da informação visual adquirida.

Em todas as estradas, mesmo naquelas que são bem estruturadas, podem acontecer mudanças de padrão que resultam de alterações climáticas ou de iluminação. O que acontece também é que a delimitação das faixas de rodagem pode ser feita de diferentes formas, linhas brancas ou bandas reflectoras, dependendo do país, região ou mesmo do tipo de rodovia. Estas variantes impõem aos sistemas que utilizam informação visual a necessidade de se adaptarem a este tipo de condicionantes.

O RALPH define-se como um sistema adaptativo, tendo como objectivo calcular a curvatura da estrada e a posição lateral do veículo relativamente à extremidade da estrada, sendo que estes parâmetros são calculados implicitamente à medida que o veículo se desloca [58, 8].

Este sistema reduz o problema de controlo lateral a 3 passos:

1. Amostragem da imagem

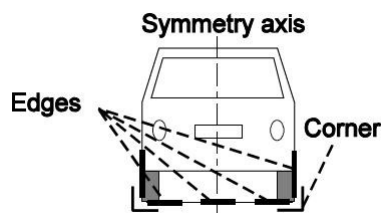


Figura 3.15: Características principais de um veículo [55].

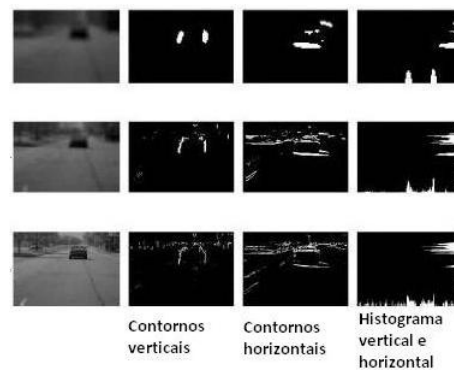


Figura 3.16: Extração de contornos verticais e horizontais [54].

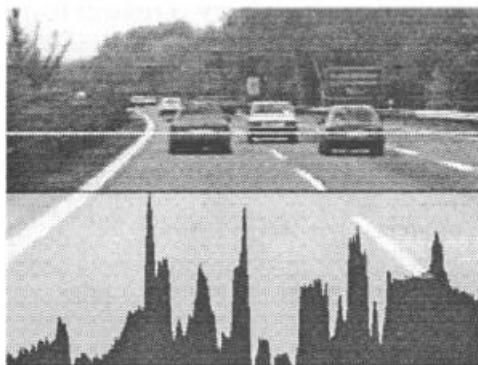


Figura 3.17: Detecção de veículos baseada em modelos [53].

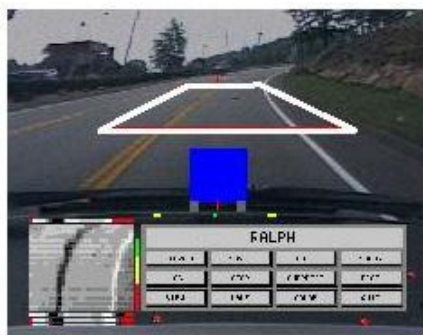
2. Determinação da curvatura da estrada
3. Determinação do posicionamento lateral do veículo

Como resultado deste procedimento é derivado um comando de actuação sobre a direcção do veículo. Este comando pode funcionar apenas como forma de aviso, ou actuar efectivamente sobre a direcção do veículo.

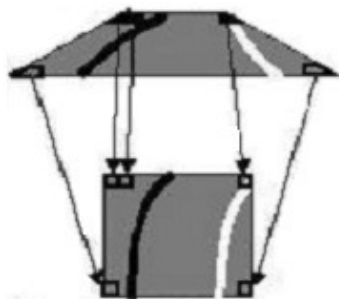
Amostragem da imagem

A aquisição de imagem é feita utilizando um câmara montada próxima do espelho retrovisor, o que permite uma aquisição tal como mostrado na figura 3.18. A imagem capturada é constituída por muitas zonas que são irrelevantes para a tarefa a executar, tais como o céu ou o próprio veículo. Estas zonas são portanto eliminadas, sendo apenas processada a zona que corresponde ao trapézio representado na figura 3.18. A altura do trapézio é dinâmica, variando de acordo com a velocidade do veículo, crescendo em direcção ao topo da imagem quando a velocidade aumenta. Em condições normais o trapézio corresponde a uma zona entre 20 e 70 metros à frente do veículo.

Uma outra característica importante deste trapézio é que a definição da sua forma horizontal é feita por forma a anular o efeito perspectiva, ou seja, se o trapézio for transformado num quadrado/rectângulo a imagem resultante é uma vista planar da zona capturada pelo próprio trapézio.



(a)



(b)

Figura 3.18: Aquisição e amostragem de imagem no RALPH: (a) aquisição de imagem utilizando uma zona trapezoidal; (b) amostragem da imagem removendo o efeito perspectiva [58].

O objectivo neste caso é criar uma imagem de baixa resolução, 30x32, em que sejam perceptíveis as características importantes da estrada, sem o efeito perspectiva associado, por forma a acelerar e simplificar o processamento seguinte (ver figura 3.18).

Determinação da curvatura da estrada

O passo anterior é crucial para a determinação da curvatura da estrada. Neste caso é utilizada uma metodologia de geração/verificação de hipóteses. O RALPH formula uma hipótese de uma possível curvatura da estrada, subtraindo esta curvatura à imagem resultante do processo anterior. Após gerar várias hipóteses de curvaturas e subtraindo estas hipóteses à imagem resultante do processo anterior, o algoritmo irá verificar qual a hipótese "vencedora", ou seja, neste caso a que gerou uma imagem com maior verticalidade.

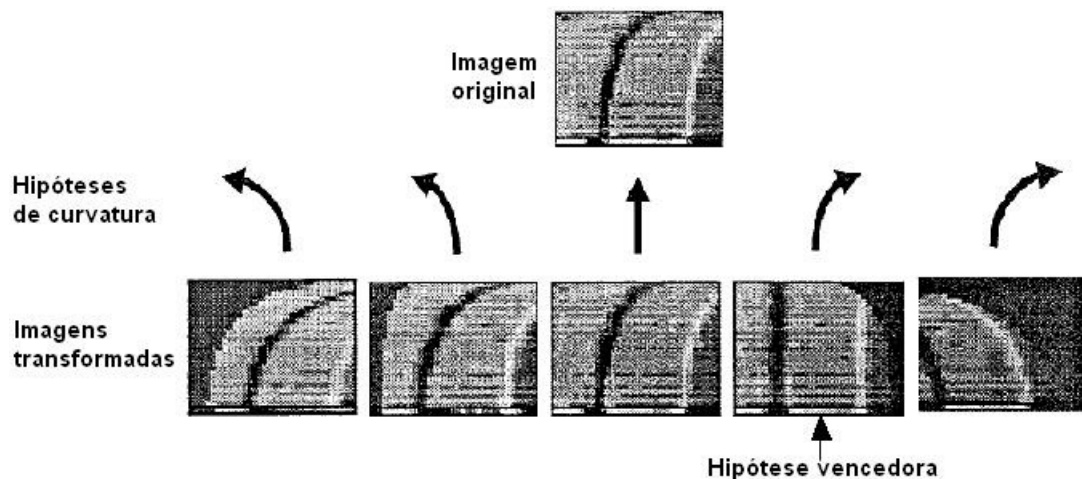


Figura 3.19: Determinação da curvatura [58].

Na figura 3.19 mostra-se a formulação de 5 hipóteses de curvatura. As linhas da imagem original (em cima) são deslocadas diferencialmente, numa tentativa de desfazer a curvatura imposta na imagem original. As linhas que se encontram próximo do topo da imagem são mais deslocadas do que as que se encontram próximo da base da mesma, o que se explica pelo efeito perspectiva: próximo do topo da imagem um único píxel representa uma região bastante maior do que um píxel que se encontra na base da mesma imagem. A distância de deslocação para cada linha, na formulação das hipóteses de curvatura, é calculada de acordo com a geometria da câmara e do nível de curvatura que se encontra nesse momento a ser testado.

Tal como pode ser visto na figura 3.19, a hipótese vencedora resultou numa transformação da imagem original em que as linhas limitadoras da faixa de rodagem estão praticamente verticais. Foi por isso encontrado um nível de curvatura que corresponde aproximadamente à curvatura da estrada.

O método de eleição da melhor hipótese de curvatura encontra-se representado na figura 3.20. É gerado um perfil de intensidade de cada imagem, representado por um gráfico/histograma de cada coluna, com a soma dos valores da intensidade de todos os píxeis desta mesma coluna. Uma linha branca ou preta pode assim ser detectada quando neste perfil de intensidade for encontrada uma descontinuidade acentuada.

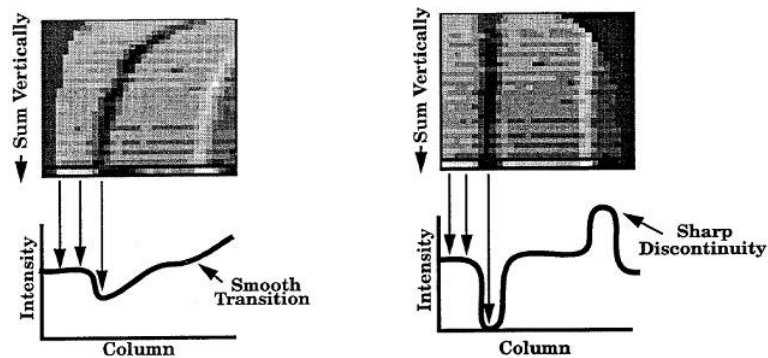


Figura 3.20: Determinação da curvatura [58].

Determinação da posição lateral do veículo

Este passo utiliza o perfil de intensidade gerado no passo anterior para estimar a deslocação lateral do veículo relativamente às faixas laterais. Inicialmente, com o veículo perfeitamente centrado na faixa de rodagem, é gerado um perfil de intensidade “modelo”. Este perfil é comparado com o perfil gerado posteriormente no passo de detecção da curvatura, permitindo assim fazer uma avaliação do desvio do veículo em relação ao centro da faixa de rodagem por comparação dos dois perfis de intensidade.

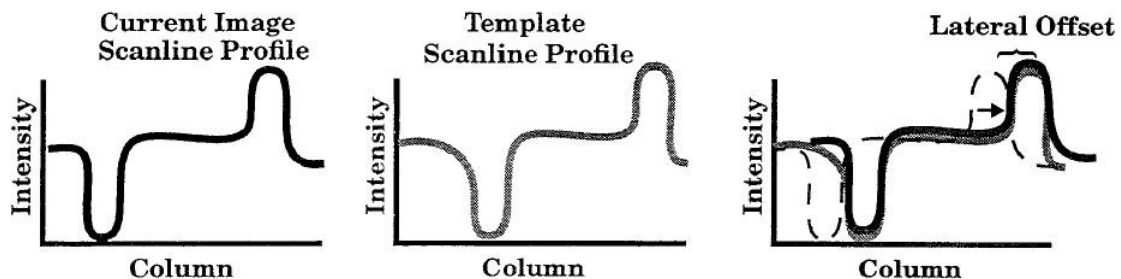


Figura 3.21: Determinação da posição lateral do veículo [58].

A figura 3.21 representa com mais detalhe o cálculo deste deslocamento lateral do veículo, em relação ao modelo inicial. Deslocando o perfil de intensidade actual para a esquerda ou direita para conseguir uma melhor correspondência com o modelo, é fácil o algoritmo conseguir calcular a distância que separa os dois, que pode posteriormente ser mapeada para a distância real, que representa o deslocamento do veículo em relação ao centro da faixa de rodagem.

Análise do desempenho do RALPH

Uma das mais importantes características do RALPH é a sua simplicidade de adaptação a novas vias, havendo apenas necessidade de calcular o seu perfil de intensidade, tal como referido nas secções anteriores. Este perfil de intensidade por ser calculado manualmente com ajuda de um condutor humano, pode ser escolhido de uma biblioteca pré-definida ou pode ainda resultar de um processo de adaptação dinâmico

3.5.4 UNSCARF

O sistema UNSCARF foi desenvolvido com o objectivo de resolver o problema da condução autónoma em ambientes não estruturados, sem linhas brancas delimitadoras e em superfícies degradadas. Utiliza uma técnica de agrupamento de píxeis com características de cor e localização semelhantes, comparando posteriormente estes grupos de píxeis com modelos de estradas pré-definidos para encontrar a melhor hipótese da localização da faixa de rodagem e consequentemente dos seus limites.

Não necessita de relacionar uma imagem com a imagem imediatamente anterior. O processo de agrupamento dos píxeis é feito de novo para cada imagem, independentemente do resultado do processo na imagem anterior.

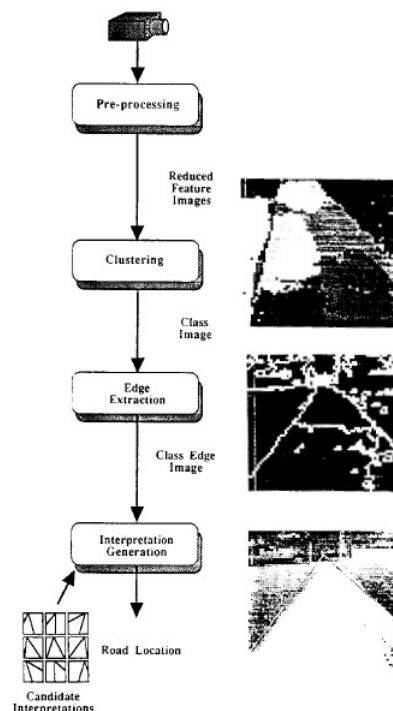


Figura 3.22: Diagrama de processamento do UNSCARF [59].

Tal como representado na figura 3.22, este método pode ser explicado dividindo o processamento em 4 etapas fundamentais:

1. Amostragem da imagem
2. Agrupamento de píxeis
3. Detecção de contornos
4. Correspondência do melhor modelo

Amostragem da imagem

Esta etapa tem como função reduzir o tamanho das imagens adquiridas e adicionar a cada imagem a informação espacial de cada píxel. A redução do tamanho da imagem é feito para diminuir o custo

de processamento, mas tem também o objectivo de reduzir o ruído imposto na imagem. A cada píxel da imagem reduzida é adicionada a sua localização espacial, formando assim uma imagem em que cada píxel x é caracterizado por um quintuplo: $x = [\text{vermelho}, \text{verde}, \text{azul}, \text{linha}, \text{coluna}]$, em que as 3 primeiras componentes dizem respeito à cor, enquanto que as duas últimas representam a localização do píxel na imagem.

Agrupamento dos píxeis

Este módulo agrupa os píxeis da imagem resultante do passo anterior para formar regiões tão homogêneas quanto possível, de acordo com um algoritmo de agrupamento explicado em detalhe em [60]. Inicialmente os píxeis são classificados arbitrariamente, ao que se segue um processo iterativo de reclassificação dos píxeis em classes, ou grupos homogêneos.

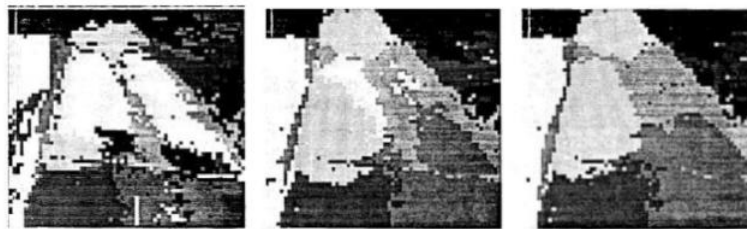


Figura 3.23: Processo iterativo de classificação dos píxeis [59].

Tal como demonstrado pela figura 3.23, o processo iterativo é executado até que a classificação se torne relativamente convergente, ou seja, que não existam mais píxeis para classificar e que se atinja um grupo de regiões relativamente homogêneas.

Deteção de contornos

Este passo do processamento utiliza a imagem resultante do passo anterior, filtrando a informação irrelevante para posteriormente poder detectar os contornos formados pelos limites dos grupos gerados. Este processo de filtragem remove os grupos mais pequenos que se consideram irrelevantes, juntando-os a outros que a eles se assemelhem para atingir uma menor granularidade na definição dos grupos.

Correspondência do melhor modelo

Esta parte final do processamento tem como tarefa interpretar os resultados dos passos anteriores e gerar um modelo hipotético de uma estrada, que mais se assemelhe a esses mesmos resultados. Este módulo compara os contornos resultantes dos processos anteriores com uma série de modelos de estrada candidatos, para verificar qual melhor se assemelha. Desta forma é escolhido um modelo, que representa em cada momento da melhor forma a estrutura da estrada.

Análise do desempenho do UNSCARF

O UNSCARF foi integrado num sistema de navegação de um veículo do projecto Navlab (ver secção 2.2.2), mostrando resultados eficientes em estradas com sombras (ver figura 3.24), pavimento degradado, limitadas por neve (ver figura 3.25) ou terra sem faixas delimitadoras.

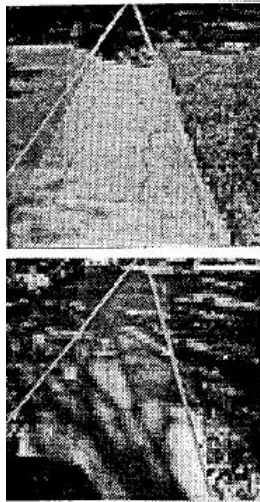


Figura 3.24: Resultados em estradas com sombras.

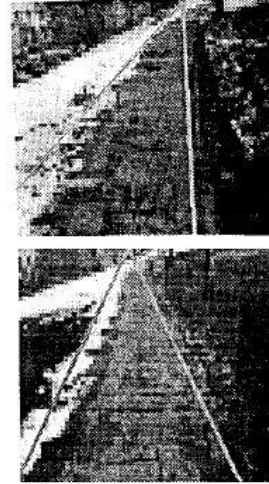


Figura 3.25: Resultados em estradas com neve.

Este sistema oferece também imunidade às mudanças bruscas de luminosidade, uma vez que o processamento é independente de frame para frame. Uma vez que a sua metodologia é baseada na análise de imagens com cor, contrariamente à generalidade dos métodos que utilizam imagens monocromáticas, o UNSCARF torna-se também bastante adaptável uma vez que consegue obter bons resultados em várias estradas com tonalidades diferentes.

O UNSCARF é um método já relativamente antigo, mas no entanto apresenta uma abordagem interessante ao problema da detecção dos limites laterais da estrada em ambiente não estruturados. Os seus resultados são bastante satisfatórios, no entanto, a elevada carga computacional que o método exige foi apontada na altura como principal ponto negativo.

3.5.5 GOLD - System for Generic Obstacle and Lane Detection

Este sistema foi criado ao abrigo do projecto EUREKA-PROMETHEUS (ver secção 2.2.1), como mais uma tentativa de melhorar a segurança rodoviária.

Baseado numa arquitectura de processamento paralelo, composto por módulos de hardware específico SIMD¹ para aumentar a capacidade de processamento, composto por duas câmaras para possibilitar visão estereoscópica, este sistema é capaz de efectuar a detecção de obstáculos genéricos, bem como reconhecimento dos limites laterais da estrada em ambientes bem estruturados, mais especificamente quando existem faixas laterais brancas no pavimento.

A câmara da esquerda é utilizada para fazer a detecção das faixas laterais, enquanto que a fusão das imagens adquiridas pelas duas câmaras, permitem fazer a detecção de obstáculos (ver figura 3.26).

Este sistema foi testado num laboratório móvel, MOBLAB (ver figura 3.27), mostrando-se capaz de o conduzir autonomamente ao longo de 3000 km de auto-estradas e vias extra-urbanas. Mais tarde este sistema foi portado para outro veículo, ARGO [61], tendo sofrido várias evoluções que o tornaram mais eficaz em certas situações [40].

Uma das tarefas iniciais deste sistema é a remoção do efeito perspectiva nas imagens capturadas, tanto da esquerda como da direita, através de uma transformação geométrica, o que permite que estas sejam analisadas como se tratasse de uma vista planar. Este mesmo efeito perspectiva associa

¹Single Instruction Multiple Data

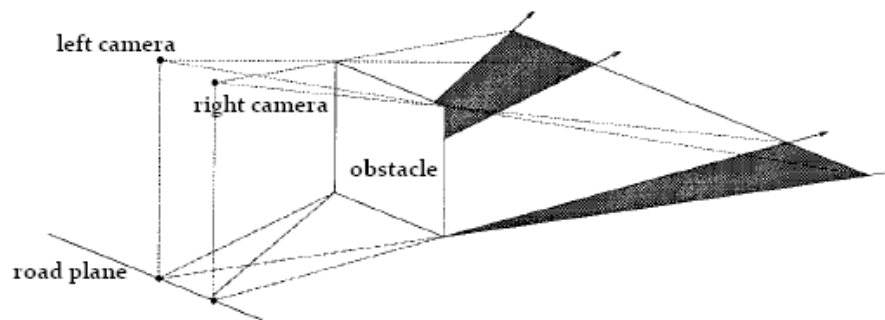


Figura 3.26: Exemplo de detecção de um obstáculo com um sistema de visão estereoscópico [40].



Figura 3.27: MOBLAB, veículo autónomo [40].

diferentes significados a diferentes píxeis na imagem, dependendo da sua posição.

O objectivo da remoção deste efeito na imagem é que cada píxel represente uma região igual na cena adquirida, permitindo uma distribuição homogénea de informação em todos os píxeis. Desta forma, o mecanismo de detecção das faixas laterais pode utilizar um modelo de correspondência com o mesmo tamanho e forma para qualquer zona da imagem.

De seguida irá ser feita uma análise aos módulos de detecção de obstáculos e dos limites laterais da estrada descrevendo assim as metodologias implementadas, referindo também por fim os pontos fracos de cada módulo.

Detecção das faixas laterais

O módulo de detecção dos limites laterais da estrada é baseado numa técnica de correspondência de padrões baseada na presença das marcas delimitadoras na própria estrada. Esta detecção é feita assumindo que as faixas laterais são representadas por linhas verticais, após a remoção do efeito perspectiva (ver figura 3.28), com um brilho muito superior à restante zona de estrada.

A figura 3.29 representa assim de uma forma simplificada o processo de detecção das faixas laterais. Inicialmente é feita a remoção do efeito perspectiva, mostrado em (b). O processo seguinte é destacar na imagem as possíveis faixas laterais, caracterizadas por um valor de brilho elevado e por serem aproximadamente verticais (c), excluindo assim as restantes zonas da imagem. De seguida é feita uma dilatação dos píxeis, tendo em conta a sua correlação vertical, tornando a imagem mais definida (d). Como passo final é feita uma binarização, para atingir ainda mais realce na imagem (e). Este processo pode ser explicado, de forma simplificada, como a definição de um valor médio de brilho a partir do qual todos os píxeis que tenham um valor inferior passam a ter brilho 0, e todos os píxeis com valor superior passam a ter brilho máximo.

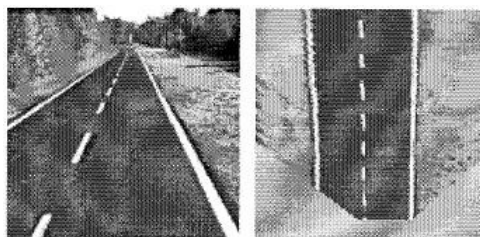


Figura 3.28: Imagem original e imagem transformada sem efeito perspectiva [40].

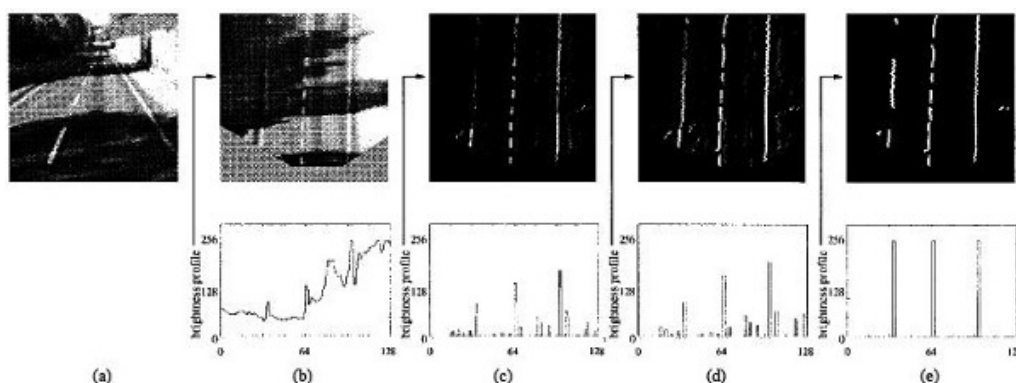


Figura 3.29: Detecção faixas laterais. (a) Original. (b) Sem o efeito perspectiva. (c) Diminuição ruído. (d) Realce (e) Binarizada [40].

Durante todo este processamento é possível verificar na figura 3.29 os histogramas correspondentes a cada imagem. No último passo do processamento, representado por (e), é fácil verificar que se identificam 3 picos no histograma que representam claramente as linhas delimitadoras da faixa de rodagem.

Este módulo de detecção das faixas laterais revela-se bastante fiável, tendo apenas algumas falhas quando a estrada não é exactamente plana, ou quando as marcas na estrada não se encontram bem visíveis.

Detecção de obstáculos

O módulo de detecção de obstáculos tenta determinar o espaço livre na estrada em frente ao veículo, em vez de tentar identificar logo quaisquer possíveis obstáculos.

Dado um modelo da estrada à frente do veículo, este sistema faz um mapeamento das imagens direita e esquerda, para tentar encontrar quaisquer diferenças que possam ser possíveis obstáculos. Mais especificamente, é feita uma subtração entre as duas imagens, esquerda e direita, após a remoção do efeito perspectiva. Possíveis obstáculos são detectados se existir um grupo suficientemente vasto de píxeis não nulos.

Tal como mostrado na figura 3.30 num caso em que o obstáculo fosse simetricamente perfeito, o resultado da subtração das duas imagens, esquerda e direita, iria resultar numa imagem com dois triângulos. A explicação da formação destes dois triângulos é a distância física entre as duas câmaras.

O processo de detecção de obstáculos é baseado na detecção destes mesmos triângulos gerados através da subtração das imagens esquerda e direita. No entanto, a detecção destes mesmos triângulos pode tornar-se uma tarefa bastante complexa devido à forma irregular, ao brilho e textura não

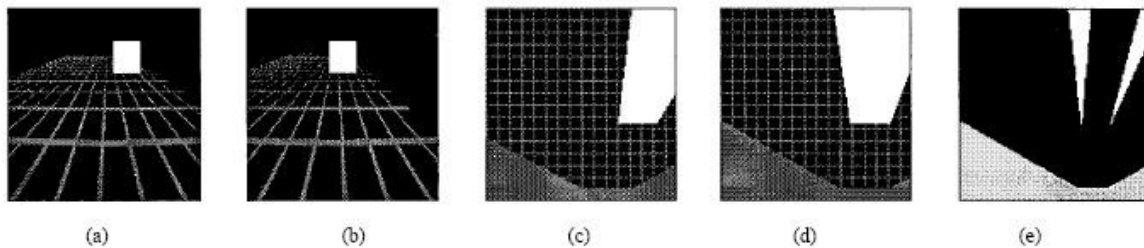


Figura 3.30: Detecção de um objecto homogéneo ideal. (a) Imagem esquerda. (b) Imagem direita. (c) Transformação da imagem esquerda. (d) Transformação da imagem direita. (e) Diferença entre (c) e (d) [40].

homogéneos nos diferentes obstáculos reais, tal como se pode verificar na figura 3.31.

Uma das principais falhas deste sistema é, quando numa determinada situação existe mais do que um veículo, ser impossível por vezes fazer a correcta detecção de cada veículo separadamente. Isto deve-se ao facto de os pares de triângulos gerados por cada veículo poder por vezes ser coincidentes, o que impossibilita diferenciar cada um, num cenário em que possam existir vários veículos próximos entre si.

A detecção da posição do obstáculo é feita através da criação de um histograma polar (ver figura 3.32), que permite assim fazer um mapeamento entre a zona da imagem onde os triângulos se encontram e o ângulo que estes formam em relação à posição do veículo.

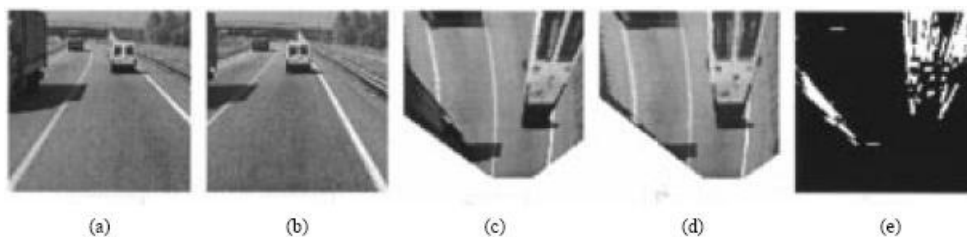


Figura 3.31: Detecção de um objecto real. (a) Imagem esquerda. (b) Imagem direita. (c) Transformação da imagem esquerda. (d) Transformação da imagem direita. (e) Diferença entre (c) e (d) [40].

Como é visível na figura 3.32 o histograma resulta claramente na formação de dois picos, que representam de uma forma grosseira os limites inferiores do obstáculo. A ligação destes mesmos dois picos representa, tal como mostrado pela figura 3.33, o limite inferior do obstáculo, existindo assim um mapeamento directo entre os picos formados pelo histograma polar e a localização do obstáculo na imagem real.

Este módulo de detecção de obstáculos, embora funcione de forma satisfatória em muitas situações, tem algumas limitações. Uma vez que é utilizada visão estereoscópica, existe uma forte dependência da calibração dos parâmetros do sistema de visão na obtenção de bons resultados.

Este sistema é também pouco fiável se os obstáculos estiverem muito distanciados do veículo, o que resulta num histograma polar muito pouco definido, dificultando a procura dos seus picos. Se o obstáculo estiver muito próximo do limiar da estrada, se estiver apenas parcialmente visível ou se o seu brilho for semelhante ao da estrada, pode também resultar numa detecção incorrecta.

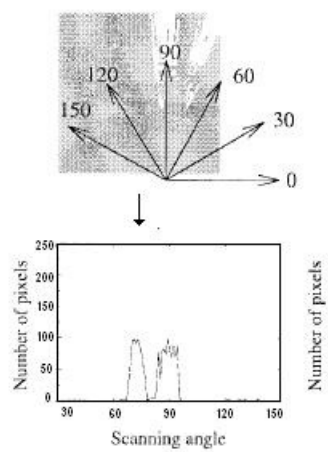


Figura 3.32: Cálculo da posição do obstáculo [40].

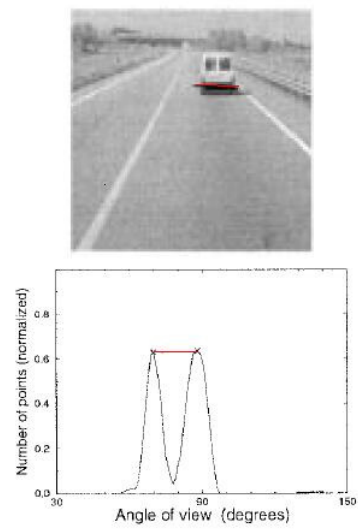


Figura 3.33: Mapeamento da posição calculada do obstáculo na imagem real [40].

Capítulo 4

Análise de imagem

Uma vez que o trabalho desenvolvido tem a finalidade de fazer um estudo de forma genérica aos métodos de processamento e análise de imagem, não existe nenhuma plataforma física de aquisição já construída para a qual o trabalho esteja especificamente a ser desenvolvido. Embora este estudo tenha também a finalidade de perceber a viabilidade da implementação de um sistema de percepção visual no HAMMERSHARK, a plataforma física do veículo não se encontra ainda concluída. Por este motivo as experiências fizeram-se em sequências de vídeo. Estas sequências tentam emular os diferentes ambientes onde será possível utilizar este tipo de sistemas.

Foram escolhidas assim duas sequências de vídeo que emulam dois ambientes estruturalmente distintos: um ambiente de pista (ver imagem do lado direito da figura 4.1) onde já se realizaram edições anteriores da prova Shell Eco-Marathon e outro que emula um ambiente real de uma auto-estrada, gravado do interior de um veículo (ver imagem do lado esquerdo da figura 4.1). Estas duas sequências de vídeo distintas tentam mostrar o funcionamento do sistema em vias estruturadas e não estruturadas, testando a sua adaptabilidade.

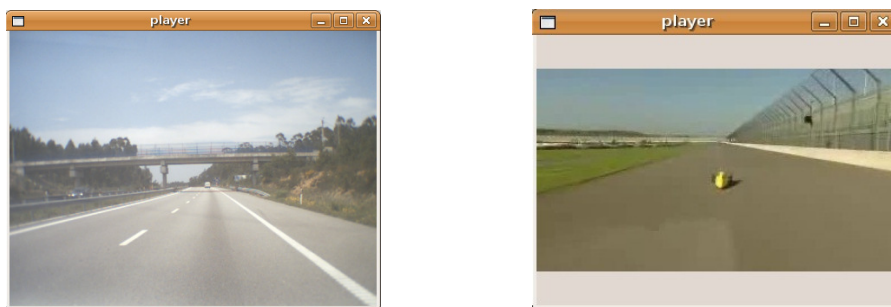


Figura 4.1: À esquerda um ambiente estruturado, capturado na A25 na zona de Ovar. À direita um ambiente não estruturado, traçado da pista Eurospeedway em Lausitz, Alemanha .

Este trabalho é então focado nas principais tarefas de um sistema de visão computadorizada para efectuar condução autónoma: detecção dos limites laterais da faixa de rodagem e detecção de obstáculos.

No entanto, para atingir estes objectivos é necessário que existam mecanismos que inicialmente sejam capazes de analisar uma imagem e fazer uma distinção entre as zonas de estrada e não estrada. Ou seja, é necessário que a zona de estrada seja inicialmente identificada para posteriormente proceder à detecção dos seus limites e à detecção dos veículos que lá circulem. Este passo tem uma importância crucial em todo o processo, uma vez que não existe um ambiente pré-definido onde o sistema se

destina a funcionar, podendo os delimitadores da faixa de rodagem serem constituídos por linhas brancas ou por outro qualquer ambiente.

Assim a abordagem feita ao problema pode ser dividida em três categorias principais: detecção da faixa de rodagem, detecção dos limites da faixa de rodagem e detecção de obstáculos. Mais adiante neste capítulo irá ser feita uma descrição mais detalhada do desenvolvimento de cada uma destas tarefas.

4.1 Biblioteca OpenCV

O OpenCV é uma biblioteca open source de visão computadorizada, desenvolvida para alto desempenho computacional e para utilização em aplicações de tempo real. O seu objectivo é disponibilizar uma infra-estrutura que ajude as pessoas a construir aplicações bastante sofisticadas em relativamente pouco tempo. Esta biblioteca contém mais de 500 funções que se espalham por muitas áreas possíveis de aplicação de visão computadorizada, entre as quais: inspecção de produtos industriais, imagem médica, segurança, interfaces de utilizador, calibração de câmaras, visão estereoscópica e robótica.

Esta biblioteca existe desde 1999, altura em que foi lançada a sua versão experimental. Desde então tem sido a base para o desenvolvimento de inúmeras aplicações nas áreas anteriormente mencionadas. É uma biblioteca compatível com Windows, Linux e Mac OS X, desenvolvida em C e C++, mas já com suporte a ser desenvolvido também para outras linguagens tais como Python, Ruby, Matlab e outras. Estas características fazem com que o OpenCV se tenha tornado uma biblioteca de referência no desenvolvimento de aplicações de visão computadorizada [45]. Foi por exemplo uma parte essencial no sistema de visão integrado no veículo que venceu a edição de 2005 do DARPA Grand Challenge (ver secção 2.3). Devido às suas características é muito apetecível para utilização na área da robótica, motivo pelo qual foi eleita para o desenvolvimento deste trabalho.

O OpenCV suporta vários formatos de imagem e vídeo assim como a aquisição directa de imagens de uma câmara, sendo ainda capaz de fazer a conversão entre vários formatos. Oferece também uma colecção de funções essenciais para processamento e manipulação de imagem que permitem executar desde os processos mais simples até processos mais complexos. Suporta os algoritmos de extracção de contornos referidos na secção 3.3.1, Canny e Sobel, que são normalmente muito utilizados em processos de análise de imagem. Oferece também suporte para as operações morfológicas referidas em 3.3.3, erosão e dilatação de píxeis. Além disto contém funções para suporte de tarefas mais complexas, tal como é o caso da aplicação de filtros de Kalman, cálculo dos vectores de optical flow ou cálculo da transformada de Hough.

O OpenCV contém também uma interface de utilizador capaz de criar janelas onde são mostradas as imagens a ser processadas, permitindo também a interacção do utilizador através do rato ou teclado, sendo possível também interagir com o processamento através do rato ou teclado,

4.2 Ferramentas de manipulação e percepção de cor no OpenCV

O formato RGB é constituído por 3 componentes, Red, Green e Blue. Cada píxel é então representado em termos cromáticos por um conjunto de três valores numéricos que no OpenCV são representados por um valor que varia entre 0 e 255. A ausência de cor é representada pelo valor 0, enquanto que a sua intensidade máxima é representada pelo valor 255, representação que é comum em outros sistemas que usam este espaço de cor.

O formato HSV é constituído igualmente por 3 componentes, matiz, saturação e brilho. No caso deste espaço de cor as componentes não têm a mesma forma de representação. As componentes

brilho e saturação são normalmente medidas em percentagem, num intervalo que varia entre 0 e 100. A componente matiz, uma vez que a sua representação assume uma forma angular, é representada num intervalo de 0 a 360. No OpenCV as componentes saturação e o brilho são representadas por um valor no intervalo de 0 a 255, enquanto que a componente matiz é representada num intervalo de 0 a 180, uma vez que não é possível representar um valor entre 0 e 360 em apenas um byte, usando assim uma representação com menor resolução para esta componente. No entanto é muito importante perceber que a matiz é representada através de uma forma angular, ou seja em módulo 180, sendo zero o valor imediatamente seguinte ao 180. A figura 4.2 mostra as cores possíveis de representar no espaço de cores HSV, no OpenCV, com as componentes saturação e brilho fixas com valores máximos. É possível verificar que as cores representadas quando a componente matiz tem valor 0 ou 180 são de facto iguais. Desta forma é necessário entender a componente matiz como sendo representada de uma forma contínua e não com valores máximos e mínimos.

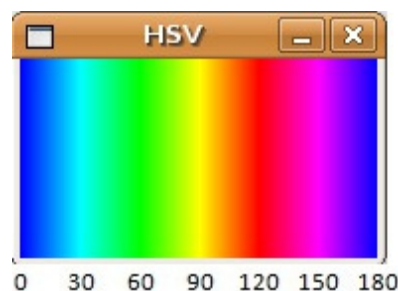


Figura 4.2: Representação da componente matiz em OpenCV.

O objectivo desta secção é apresentar o desenvolvimento de pequenas ferramentas de suporte à interacção com uma imagem em formato HSV, de forma a melhor perceber a composição deste espaço de cor e a implicação de cada uma das componentes numa imagem neste espaço de cor.

O sistema de percepção visual humano permite-nos estimar mais facilmente os valores de um píxel no espaço HSV do que no espaço RGB. No entanto esta é uma tarefa que se torna complicada em ambientes reais. Assim por forma a perceber melhor os valores de cor presentes nas sequências de vídeo analisadas, foi criada uma pequena aplicação capaz de devolver os valores das componentes da cor de um qualquer píxel da imagem, que é escolhido com um click do rato. Se forem escolhidos múltiplos píxeis, os valores mostrados são a gama de valores HSV em que se encontram os píxeis escolhidos. Ou seja, se escolhermos vários píxeis, o que será mostrado serão os valores máximos e mínimos de cada componente, H, S e V. De notar que a componente matiz é a única em que o valor mínimo pode ser superior ao máximo. Neste caso é escolhido o menor intervalo possível para representar todos os valores escolhidos.

Na primeira imagem da figura 4.3 foi seleccionado apenas um ponto. Tal como é possível verificar, olhando para os valores representados pelas trackbars para cada componente, a composição deste ponto no formato HSV é (90,13,156) para as componentes matiz, saturação e brilho.

Na segunda imagem da figura 4.3 foi feita uma selecção de vários pontos tendo o cuidado de escolher apenas os que situam na faixa de rodagem, por forma a tentar perceber se seria possível identificar características comuns nestes mesmos pontos. A olho nu estes pontos parecem realmente ter características muito semelhantes, o que se verifica nas componentes da saturação e do brilho que apresentam intervalos bastante pequenos, o que significa que estas componentes “convergem” para um determinado valor. O que se passa na componente matiz é que os pontos seleccionados resultam num intervalo de valores nesta componente que abrange uma grande percentagem de todo o intervalo possível. A componente matiz é representável por um intervalo com amplitude máxima de 180, como



Figura 4.3: Ferramenta de selecção de pontos para visualizar a sua composição cromática.

já tinha sido referido, e neste exemplo a gama de valores da componente matiz correspondente aos pontos seleccionados resultou num intervalo com amplitude de 114.

Nas componentes brilho e saturação o intervalo encontrado é relativamente bastante mais pequeno, quando comparado com a componente matiz. Uma possível justificação para este facto é que a zona de onde os pontos foram extraídos, a faixa de rodagem, é constituída maioritariamente por cinzento, e se tivermos em atenção o espaço de cores HSV é possível verificar que esta cor é constituída por valores baixos de saturação, ou seja, pouca quantidade de cor. Este facto faz com que as variações na componente matiz não tenham grande impacto na representação da cor cinzento, sendo possível representar várias tonalidades de cinzento semelhantes, com valores distintos na componente matiz.

A figura 4.4 representa uma pequena aplicação que permite a partir de uma determinada imagem representada em HSV eliminar certas zonas que não correspondam às características definidas pelo utilizador. É possível então definir intervalos válidos, para cada uma das componentes HSV, segundo os quais serão eliminadas da imagem todas as zonas que não correspondam a estas características. Ou seja, a imagem resultante irá ser igual à imagem original nas zonas que tenham os valores das componentes representadas dentro dos intervalos definidos.

No caso da imagem do lado esquerdo da figura 4.4 não existem zonas eliminadas porque os intervalos definidos em todas as componentes têm amplitude máxima, ou seja, abrangem todos os valores. Na imagem do lado direito o intervalo da componente matiz foi definido por forma a serem apenas representadas zonas que tenham a sua cor semelhante ao azul. O azul no OpenCV é representado aproximadamente pelo intervalo $[165 - 15]$, tendo mais uma vez em conta a representação circular deste espaço de cor. Como resultado podemos ver que uma das zonas mais identificáveis é o sinal de trânsito que se encontra na parte lateral direita da imagem, uma vez que a sua cor dominante é o azul. Na imagem do meio o intervalo escolhido na componente do brilho resulta no isolamento apenas das zonas com um nível de brilho superior, o que neste caso resulta na eliminação de todos os pixels que não pertencem ao céu, uma vez ser esta a zona da imagem que é composta por um brilho mais forte.

4.3 Impacto das sombras

Esta secção dedica-se ao estudo e compreensão do impacto das sombras nas características cromáticas das imagens de estradas.

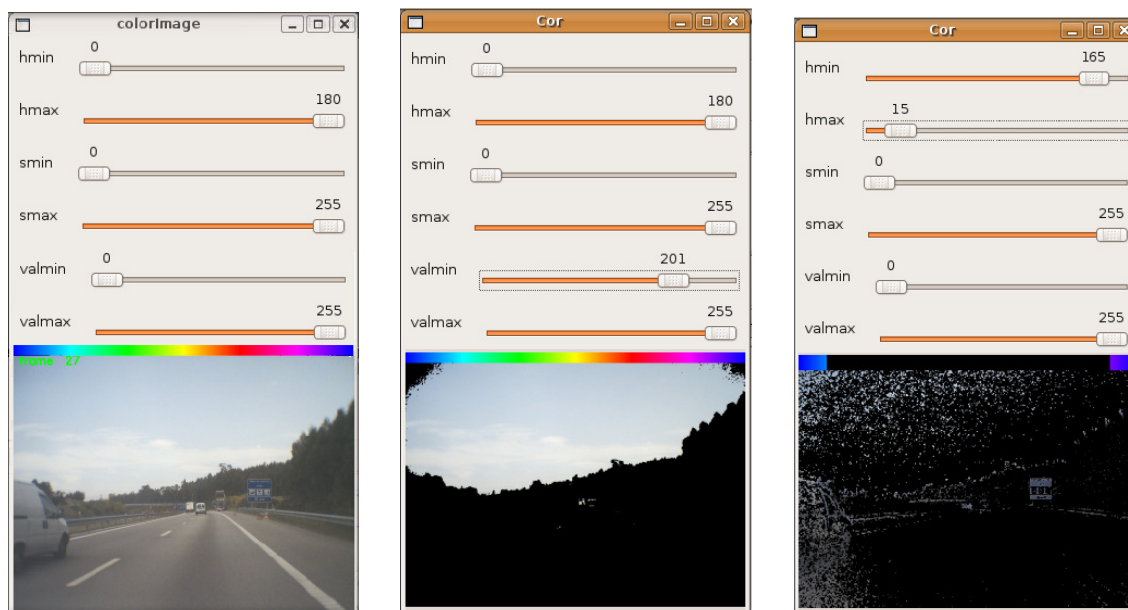


Figura 4.4: Da esquerda para a direita: a imagem original, segmentação de zonas com valor elevado de brilho e segmentação de zonas azuis.

Podemos definir sombra, de uma forma simples, como uma região em que a luminância é inferior à das regiões que a circundam, por efeito da obstrução da fonte de luz primária. No entanto, qual será o impacto de uma sombra numa qualquer região, em termos das suas características cromáticas? Para tentar encontrar resposta a esta pergunta, uma pequena experiência foi feita, representada pelas imagens na figura 4.5.

Na imagem da estrada foi escolhido um conjunto de pontos, que intersectam zonas com sombra e sem sombra. Nos gráficos representados na figura 4.5 estão representados os valores das componentes de todos os pontos, em dois espaços de cor distintos, tendo os pontos sido colocados por ordem crescente de Y. Existem 3 zonas de análise distintas: a zona inicial que é composta por 5 pontos e não é afectada por qualquer sombra, a zona intermédia que é composta por 4 pontos e é afectada pela sombra de um viaduto e a zona final que é composta por 3 pontos e também não é afectada por sombra. Estas 3 zonas estão demarcadas nos gráficos na figura 4.5.

No espaço de cores HSV, representado pelo primeiro gráfico, pode-se constatar que:

- A alteração na componente saturação é praticamente nula entre as diferentes zonas.
- Na componente brilho é bem visível que na zona com sombra existe uma queda brusca, para novamente voltar a subir bruscamente na transição para a zona sem sombra.
- A componente matiz é a de mais difícil análise. O que acontece é que nas zonas sem sombra mantêm um valor relativamente constante e na zona com sombra atinge valores máximos e mínimos (o valor máximo da componente matiz é 180, na sua representação no OpenCV)

A saturação representa a quantidade de cor e uma vez que estas zonas são representadas maioritariamente por cinzento, a saturação irá ser sempre baixa, quer se trate de uma zona com ou sem sombra. O brilho é a componente onde a alteração entre zonas de sombra e não sombra é mais visível. A componente matiz é a de mais difícil compreensão pela sua aparente variação no mesmo tipo de

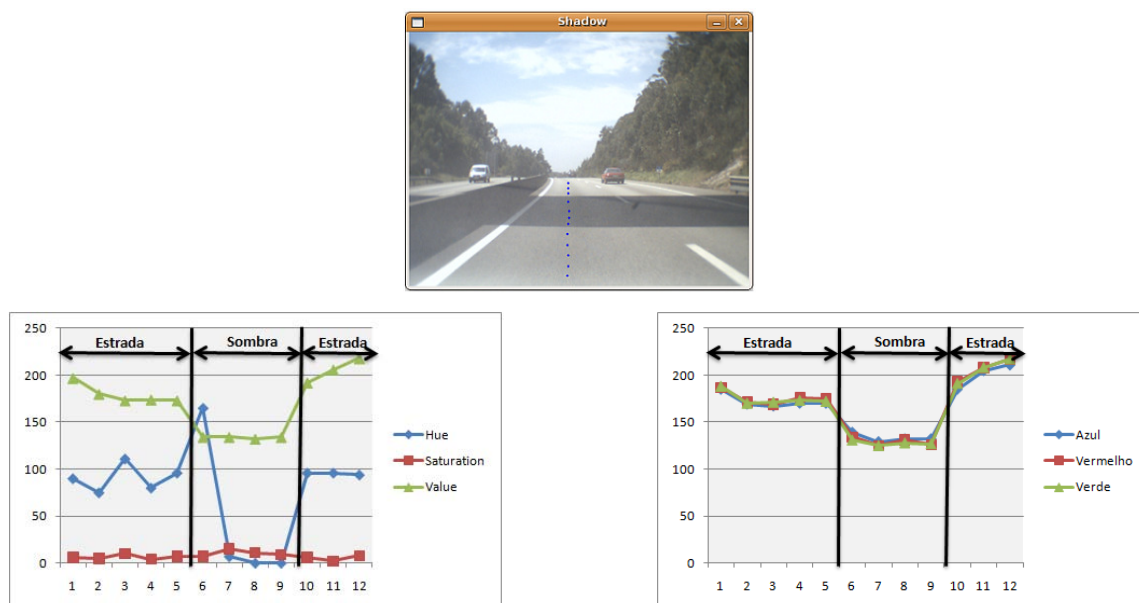


Figura 4.5: Em cima a imagem original de onde foram escolhidas várias amostras. Em baixo, da esquerda para a direita, os gráficos dos espaços de cor HSV e RGB, respectivos às amostras retiradas da imagem de cima.

zonas. Na zona de sombra esta componente atinge valores máximos e mínimos, mas para compreender estes valores é muito importante ter em atenção que a matiz é representada de forma circular. Desta forma, ao contrário do que pode parecer numa primeira análise, os valores da componente matiz na zona de sombra estão de facto muito próximos. A variação desta componente nas zonas não afectadas pela sombra pode ser explicada através da análise do cone que normalmente é utilizado para representar o espaço de cor HSV. Se analisarmos a representação deste espaço de cor podemos verificar que os tons de cinzento se encontram no eixo vertical interior deste mesmo cone, sendo definidos maioritariamente por valores baixos de saturação. Uma vez que a saturação se refere à quantidade de cor, um valor baixo de saturação faz com que variações na componente matiz sejam por vezes difíceis de distinguir visualmente. Ou seja, é possível criar tons de cinzento muito semelhantes com valores na componente matiz diferentes.

No espaço de cores RGB, representado pelo segundo gráfico, podemos verificar que as três componentes têm valores muito semelhantes em todas as zonas, notando-se a sua descida na zona intermédia de sombra. Isto deve-se à própria composição do espaço de cores RGB. Para representar o preto neste espaço de cores todas as componentes tem de ter valor nulo, enquanto que para representar o branco todas as componentes tem que ter valor máximo. O cinzento é então representado quando todas as componentes são não nulas e muito semelhantes entre si, sendo desta forma possível representar várias tonalidades de cinzento. A transição que podemos verificar no gráfico do lado direito da figura 4.5 é consequência disto mesmo. A zona que não é afectada por sombra é cinzenta tal como a zona afectada por sombra, só que na zona de sombra todas as componentes são mais baixas, aproximando este mesmo cinzento da cor preto.

Como conclusão desta experiência pode-se dizer que as sombras no espaço de cores HSV são caracterizadas principalmente por uma descida brusca na componente de brilho. No espaço de cores RGB, a transição de uma zona sem sombra para uma zona com sombra pode ser definida como uma descida, não muito acentuada, em todas as 3 componentes.

Existe no entanto outro aspecto importante referido na literatura relativo à detecção de sombras,

que irá ser verificado de seguida. Este trabalho é direccionado para utilização em ambientes externos onde a fonte de iluminação é o sol, contando no entanto com a contribuição do céu que caso esteja nublado irá tornar a iluminação mais difusa. As sombras, uma vez que são apenas iluminadas a partir de uma fonte indirecta (maioritariamente o céu, por reflexão), tendem a ser mais azuis que as restantes zonas circundantes que são também iluminadas pelo sol, uma vez que o céu se caracteriza pela sua cor azul.

Por exemplo, digamos que o triplo (pr,pg,pb) representa o valor dos píxeis em zonas que estrada e o triplo (pR,pG,pB) representa os píxeis em zonas com sombras, no espaço de cor RGB. A relação entre pB/pb será supostamente maior que a relação entre pg/pG ou pr/pR [62].

Como forma de melhor perceber esta afirmação descrita na literatura realizou-se mais uma experiência. As imagens representadas na figura 4.6 mostram um exemplo de uma sombra causada por um viaduto e de várias sombras causadas pela presença de veículos. Vários pontos foram retirados destas imagens, em zonas de sombra e em zonas não afectadas por nenhuma sombra.



Figura 4.6: Recolha de amostras de pontos em zonas de sombra (pontos a vermelho) e sem sombra (pontos a verde).

Estes pontos foram usados para construir os gráficos representados na figura 4.7

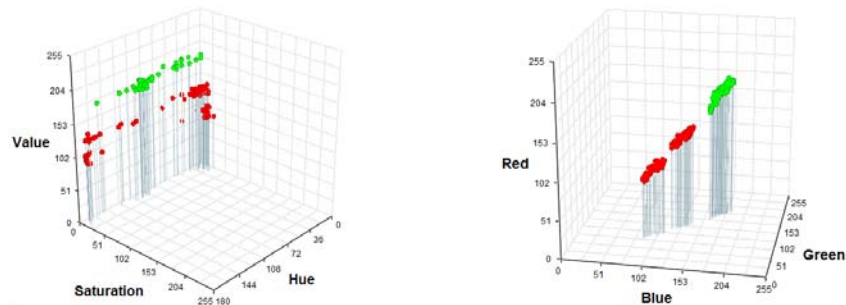


Figura 4.7: Gráfico das amostras retiradas das imagens na figura 4.6. Os pontos a verde representam os pontos em zonas sem sombra, os pontos a vermelho representam os pontos em zonas de sombra.

Nos dois gráficos da figura 4.7 podemos verificar que existem pontos vermelhos e verdes: os pontos a vermelho representam pontos retirados em zonas de sombra e os pontos a verde em zonas sem sombra, tal como mostrado na figura 4.6.

Quanto ao espaço de cores HSV, representado no primeiro gráfico da figura 4.7, podemos verificar que os pontos diferem entre si na componente de brilho, tendo os pontos retirados das zonas com sombra um valor bastante mais baixo que os restantes. Quanto à componente saturação podemos referir que é baixa nos pontos com e sem sombra, reafirmando o que já tinha sido provado na experiência anterior. A componente matiz revela nestes gráficos características que já eram visíveis nos gráficos da experiência anterior, embora não se tenha feito uma análise nesse sentido porque nesta experiên-

cia estas características são ainda mais notórias. Tal como tinha sido referido, existem referências na literatura que afirmam que as zonas de sombra têm características mais azuis do que as restantes. Este facto pode realmente verificar-se, se olharmos para a distribuição dos pontos na componente matiz: os pontos com sombra encontram-se mais próximos das extremidades e os pontos sem sombra aglomeram-se na zona intermédia.

A componente matiz revela nestes gráficos características que já eram visíveis nos gráficos da experiência anterior, embora não se tenha feito uma análise nesse sentido porque nesta experiência estas características são ainda mais notórias.

Na figura 4.8 está representado o mesmo gráfico que na primeira imagem da figura 4.7, apenas com outro ponto de vista e uma banda exemplificativa do espectro da componente matiz.

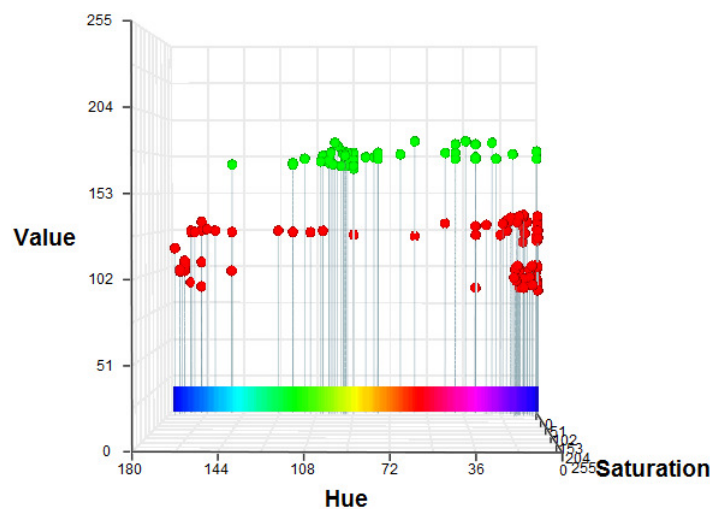


Figura 4.8: Gráfico relativo às amostras recolhidas nas imagens da figura 4.6.

Tal como é possível então verificar, os pontos vermelhos aglomeram-se mais nas extremidades onde se encontram as zonas azuis na componente matiz, enquanto que os pontos verdes incidem com maior expressão na zona intermédia. Verifica-se então que as zonas de sombra possuem características mais azuis que as restantes.

No entanto, a componente matiz continua bastante irregular, uma vez que existe ainda uma quantidade considerável de pontos que não obedece a esta regra. Isto deve-se ao facto de se tratar de pontos que possuem um valor de saturação muito baixo e por isso resulta em variações no espectro da componente matiz bastante acentuadas.

No gráfico representado pela segunda imagem da figura 4.7, onde se encontra representado o espaço de cores RGB, não se consegue verificar a mesma propriedade das zonas com sombra. As variações das zonas sem sombra para zonas com sombra distinguem-se por valores mais baixos em todas as componentes, sendo que estas componentes têm valores muito semelhantes, quer numa ou noutra situação. No entanto não é possível verificar neste espaço de cores a mesma característica verificada no espaço de cores HSV. Os pontos que dizem respeito a zonas de sombra não revelam portanto características mais azuis, uma vez que a componente azul de todos os pontos tem uma relação semelhante com as outras componentes quer seja em zonas de sombra ou não sombra.

4.4 Detecção da faixa de rodagem

Apresentação do problema

Um dos objectivos fundamentais da condução autónoma é o de manter o veículo a circular dentro dos limites da faixa de rodagem. Para isto é necessário que ele, o veículo, identifique essa faixa de rodagem. Em ambientes reais essa tarefa pode ser fortemente dificultada devido a diversos factores. Mesmo em estradas bem estruturadas, como é o caso de uma auto-estrada, o piso não tem um padrão uniforme. Devido a zonas reparadas, passagem em viadutos ou outras situações anormais, há segmentos de estrada que têm um piso com padrão diferente. A luz, ou a ausência dela, provoca alteração significativa na forma como o piso é percebido pelo olho humano ou por um sistema de visão computadorizada.

Pretendeu-se nesta componente do trabalho analisar abordagens de detecção de faixas de rodagem que minimizassem esta diferença, potenciando a obtenção de um algoritmo que faça uma boa detecção. A ideia é a de, partindo de uma imagem original, obter uma imagem binarizada onde a faixa de rodagem seja pintada a branco e o resto a preto. Este processo é fundamental para que o algoritmo de detecção tenha alguma eficiência, uma vez que por vezes as imagens adquiridas não têm as suas características muito definidas ou visíveis como seria desejável, sendo necessário então algum processamento anterior para aumentar o realce destas mesmas características. Uma vez que o objectivo é alcançar um sistema que consiga responder em situações em que a estrada se apresenta de uma forma estruturada ou não estruturada, este processo assume ainda uma maior importância.

As características que se pretendem realçar são portanto as linhas brancas laterais, caso existam, ou os limites da faixa de rodagem, desconhecendo neste caso o tipo de terreno ou ambiente que poderemos encontrar a fazer limites com esta mesma faixa.

Definição da estratégia

A estratégia seguida foi a de usar sensores padrão, colocados sobre a imagem em regiões com alta probabilidade de ser estrada. A análise da cor dos píxeis cobertos pelos sensores define intervalos de cor no espaço HSV usados para segmentar a imagem em faixa de rodagem e resto. Ou seja, este processo define-se então por uma tentativa de identificar as características da zona de estrada e eliminar da imagem todas as regiões com características diferentes. À medida que o carro se move os sensores cobrem novas zonas da estrada, criando desejavelmente um mecanismo adaptativo de medição do padrão do piso da faixa de rodagem.

Há no entanto vários aspectos a definir de modo a que o mecanismo possa ser bem sucedido. Por exemplo, quantos sensores usar? A figura 4.9 mostra a imagem de uma estrada com dois sensores padrão desenhados e a segmentação dessa mesma imagem com base em intervalos de cor definidos por esses sensores. Os dois sensores definem um padrão válido da estrada muito restrito, uma vez que são apenas dois e estão muito próximos. Claramente, a segmentação é pobre, deixando de fora muitas zonas de estrada, embora não tenha incluído nada que não seja faixa de rodagem ou berma.

Aumentando o número de sensores ou aumentando a sua dimensão alargar-se-á a zona coberta, mas também traz desvantagens: aumenta-se o tempo de processamento e aumenta-se a probabilidade de que um sensor fique fora da zona de interesse.

Processamento de cada sensor de padrão

Cada sensor de padrão define então uma gama ou intervalo, em cada uma das componentes, que irá passar a ser aceite como região válida de estrada. Como exemplo temos a figura 4.9 que demonstra

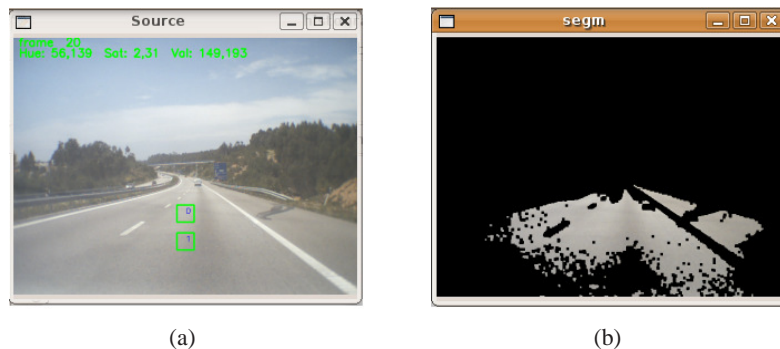


Figura 4.9: Segmentação da faixa de rodagem com base em 2 sensores padrão. (a) Imagem original com os dois sensores marcados. (b) Segmentação da imagem de acordo com os sensores padrão.

a segmentação a ser feita apenas com dois sensores padrão.

As regiões abrangidas pelos sensores padrão são avaliadas de forma a obter informação sobre a cor de todos os píxeis que nela se encontram. São definidos então três intervalos, um para cada componente, utilizando o mínimo e o máximo valores de cada componente dos píxeis das regiões abrangidas pelos sensores.

Como exemplo é possível verificar os intervalos criados pelos dois sensores padrão representados na figura 4.9:

- | | |
|-------------------|-------------------|
| • H : [81 - 127] | • H : [61 - 153] |
| • S : [4 - 37] | • S : [2 - 32] |
| • V : [158 - 202] | • V : [144 - 190] |

No entanto, estas regiões não são automaticamente aceites como definindo um padrão da faixa de rodagem. Adiante se verá que algumas das regiões serão classificadas como não válidas, não sendo nesse caso consideradas para a definição do padrão da estrada. Por exemplo, a posição de um sensor pode coincidir com a de um obstáculo, não tendo neste caso nada a ver com a faixa de rodagem.

Conjugação de sensores

Distribuindo sensores ao longo da faixa de rodagem pode-se melhorar a detecção do padrão que caracteriza a faixa de rodagem. Levanta-se então a questão de como conjugar cada intervalo de cor definido por cada sensor de forma a obter um intervalo global do padrão da estrada.

Tendo como exemplo novamente a figura 4.9 e os intervalos resultantes dos dois sensores, os sensores são conjugados através da união dos intervalos de cada um, por forma a definir então o padrão global da estrada. Neste caso específico, a conjugação das gamas de valores definidas pelos sensores da figura 4.9 resultam num padrão global com os seguintes valores:

- H : [61 - 153]
- S : [2 - 37]
- V : [144 - 202]

O que significa que todas as regiões da imagem que se encontrem dentro desta gama serão consideradas como zonas que pertencem à estrada. Como resultado é possível verificar na imagem da direita da figura 4.9 a segmentação efectuada, sendo eliminadas da imagem todas as zonas que não correspondem ao padrão que foi encontrado pelos sensores. À medida que são adicionadas mais sensores, a gama válida vai sendo aumentada e como consequência vai aumentando a região da estrada que é considerada como válida.

Validação dos sensores

Foi referido anteriormente que nem todos os sensores são considerados válidos enquanto contribuintes para a definição do padrão de cor da faixa de rodagem. Na realidade, os sensores são avaliados e classificados como válidos ou não válidos. Apenas os válidos são usados.

Para melhor perceber os valores de cor das zonas abrangidas pelos sensores, foram criados mecanismos que permitem criar um histograma relativo a cada sensor, para ser possível visionar quais os valores de cada componente de forma mais intuitiva. É também possível agrupar vários histogramas num só, relativos a vários sensores, por forma a perceber então qual será o perfil dominante da estrada.

A figura 4.10 mostra os histogramas de cada componente de cor, H, S e V, resultantes da análise das regiões abrangidas pelos sensores representados na imagem da esquerda da figura 4.9.

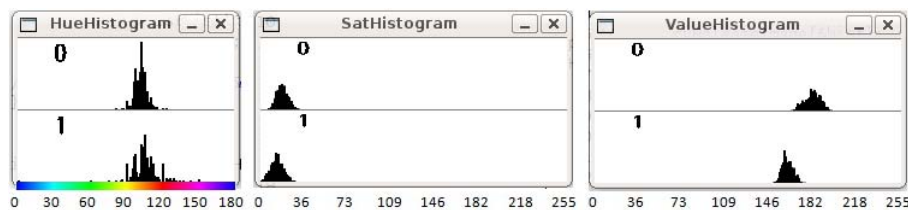


Figura 4.10: Histogramas resultantes da análise cromática dos sensores da figura 4.9.

É possível o uso de vários sensores padrão, em vez de apenas dois como demonstrado anteriormente, para gerar vários histogramas. Desta forma é possível identificar com maior facilidade as características de cor dominantes nas regiões analisadas.

Para mais fácil visualização estão impressos nas imagens anteriormente mencionadas um número que identifica unicamente cada sensor, número este que está também impresso nos histogramas para ser possível fazer a correspondência entre cada sensor e o respectivo histograma.

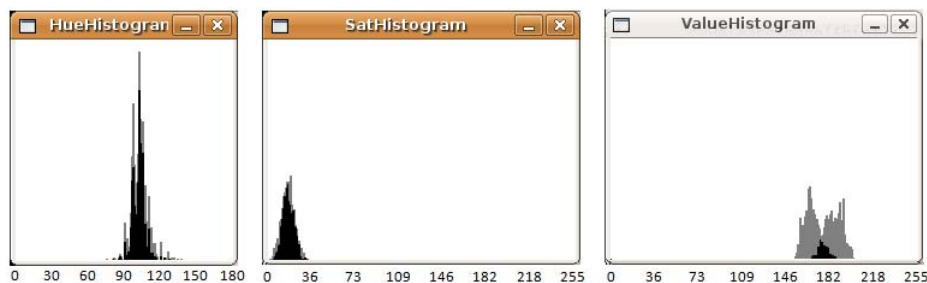


Figura 4.11: Histogramas resultantes da análise cromática global de todos os sensores da figura 4.9, em cada componente.

As figuras 4.10 e 4.11 representam assim de forma mais “amigável” as características de cor das regiões abrangidas pelos sensores da figura 4.9. Na figura 4.11 são representados apenas 3 histogramas, um para cada componente HSV. Estes histogramas representam assim uma análise global de

uma determinada componente, tendo em conta todas as regiões analisadas. São construídos sobrepondo todos os histogramas de cada componente de todas as regiões abrangidas pelos sensores. As zonas mais intensas representam assim zonas em que se verifica uma maior sobreposição dos valores, revelando-se por isso uma zona dominante.

Tal como já tinha sido referido anteriormente, para aumentar o desempenho deste processo de segmentação é necessário aumentar o número de sensores, que se posicionem em zonas dispersas da estrada. No entanto, aumentando o número de sensores irá aumentar a probabilidade de que algum deles se situe fora da faixa de rodagem ou se situe nas zonas de faixas brancas, caso elas existam, causando uma má segmentação da estrada.

O que é mostrado pela figura 4.12 é o facto de por vezes um sensor capturar uma região pertencente a uma linha branca, o que faz com que este padrão da linha branca seja aceite como região válida da estrada, não sendo assim segmentado como seria desejável. Como exemplo podemos verificar que nas regiões inferior esquerda e direita da imagem do lado direito da figura 4.12 existem partes das linhas delimitadoras que são aceites como regiões válidas não sendo assim segmentadas.



Figura 4.12: Exemplo de má segmentação devido ao incorrecto posicionamento dos sensores.

Para resolver este problema foi implementado um mecanismo que classifica um sensor como válido ou inválido, tendo em conta a sua similaridade com os restantes sensores. Um sensor que é considerado inválido é ignorado no momento de construir o padrão global da estrada.

Para este mecanismo de validação dos sensores, foi utilizada apenas a informação de uma componente, o brilho. O brilho é a componente onde é possível obter mais regularidade e é onde certamente se denota mais variância caso exista algum sensor que esteja situado numa região que se encontre fora do padrão da estrada.

O processo de validação resume-se aos seguintes passos:

1. Cálculo da média de cada sensor na componente brilho.
2. Cálculo da média das médias de todos os sensores na componente brilho.
3. Calcular o desvio padrão associado à média calculada no passo anterior.
4. Verificar individualmente cada sensor, verificando se a sua média na componente brilho se encontra dentro do intervalo definido pelo desvio padrão. Caso contrário, a região é considerada inválida.
5. Verificar individualmente cada sensor verificando se o seu próprio desvio padrão é demasiado alto, na componente brilho. Caso se verifique, a região é considerada inválida.

No passo 4 a verificação feita tem como objectivo eliminar sensores que se encontrem completamente fora da estrada, e que por isso revelam um valor na componente brilho bastante diferente dos restantes sensores. O passo 5 tem como objectivo invalidar sensores que se encontrem em zonas de transição, por exemplo faixas brancas ou limites da estrada, e que por isso revelam um desvio padrão fora do comum.

Podemos ver por exemplo nas imagens do lado esquerdo da figura 4.13 que alguns dos sensores se encontram a vermelho. Estes sensores foram considerados inválidos e portanto não são considerados para construir o padrão global da estrada. A razão pela qual cada sensor foi eliminado está impressa (de forma abreviada) próximo de cada sensor inválido, para facilitar a compreensão.

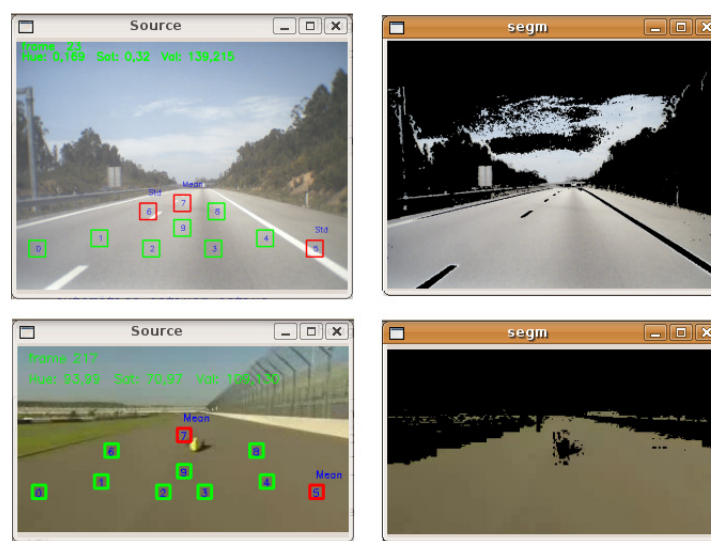


Figura 4.13: Segmentação da imagem usando sensores padrão.

As figuras 4.14 e 4.15 representam os valores da componente brilho de cada sensor das imagens do lado esquerdo das figuras 4.13 respectivamente. Desta forma é mais fácil perceber o motivo pelo qual os sensores são considerados não válidos. Tendo como exemplo a imagem superior esquerda da figura 4.13 podemos verificar que foram considerados como inválidos os sensores 5, 6 e 7. Os sensores 5 e 6 como é possível verificar estão em zonas de transição entre estrada e faixas brancas. Olhando para a figura 4.14 podemos verificar que os histogramas correspondentes a estes mesmos sensores apresentam uma certa dispersão nos valores da componente brilho. Estes sensores são assinaladas como inválidos e é impresso um "Std" imediatamente por cima, como abreviatura de Standard deviation, na imagem superior esquerda da figura 4.13. Ou seja, o desvio padrão destes sensores (35 e 19) na componente brilho é anormalmente superior às restantes (normalmente inferiores a 7).

O sensor 7 está numa zona que apanha uma pequena transição de faixa branca, mas está também numa zona que é ligeiramente mais brilhante que as restantes zonas de estrada. Embora o seu desvio padrão seja já um pouco elevado (11), a razão pela qual esta região é eliminada é o facto da média desta mesma região estar fora do intervalo admissível calculado através da média e desvio padrão de todas as regiões. Este facto é possível de verificação tendo em atenção a figura 4.14, onde os valores dominantes da zona 7 se apresentam um pouco mais para a direita que os restantes.

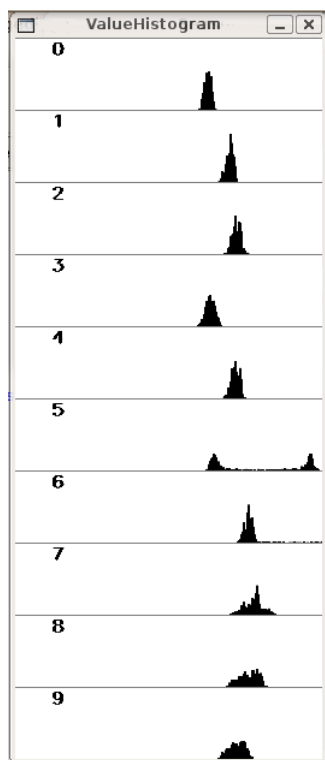


Figura 4.14: Histogramas das regiões de análise na imagem superior esquerda da figura 4.13.

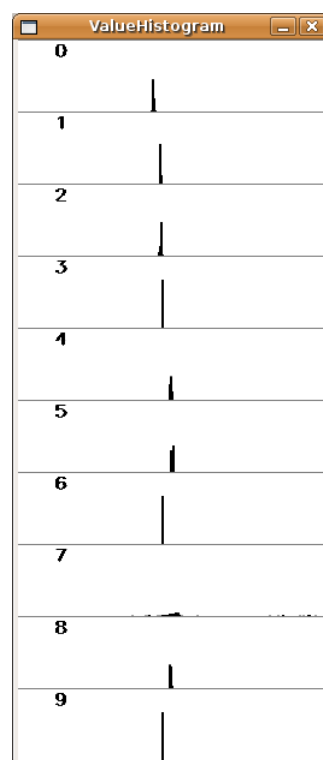


Figura 4.15: Histogramas das regiões de análise na imagem inferior esquerda da figura 4.13.

Melhoramento da detecção

As imagens da figura 4.13 mostram várias zonas intermédias que não são correctamente segmentadas. Como exemplo, na imagem superior direita da figura 4.13 existe do lado esquerdo um poste que faz parte da sinalização local que não é correctamente segmentado, uma vez que a sua cor se aproxima da cor da faixa de rodagem. A figura 4.13 apresenta também várias zonas que não são correctamente segmentadas, principalmente a zona que se aproxima da linha do horizonte.

Estas zonas podiam ser eliminadas correctamente se o processo de segmentação tivesse uma maior granularidade. Ou seja, o que é necessário fazer neste caso é a aplicação de operações morfológicas que possam eliminar estas zonas de pequena dimensão que se caracterizam por estar no meio das zonas que se encontram correctamente segmentadas. As imagens da figura 4.16 resultam da aplicação de um processo de erosão (duas iterações) seguido de um processo de dilatação (uma iteração) às imagens do lado direito da figura 4.13 respectivamente.

A dilatação efectuada permite que as zonas correctamente segmentadas se expandam, eliminando zonas que tenham sido incorrectamente segmentadas. A erosão tem como objectivo refinar o resultado, tornando as zonas de contraste mais contínuas.

Após a segmentação inicial da imagem é feita uma binarização do seu resultado, ou seja, tudo o que foi considerado zona de não estrada está a preto e tudo o que resta passa a branco. Este processo previne por vezes que existam pequenas marcas ou diferenças de cor na estrada, que podem ser entendidos como limite, sendo que desta forma teremos uma imagem mais simples e fácil de analisar. A figura 4.17 é o resultado da binarização da imagem da esquerda da figura 4.16.



Figura 4.16: Imagens correctamente segmentadas após a aplicação de operações morfológicas.

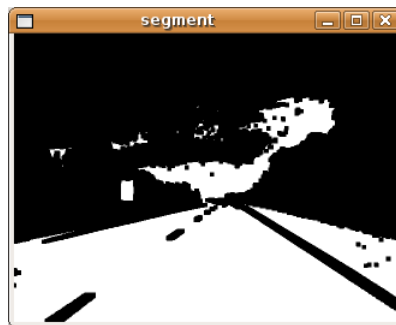


Figura 4.17: Binarização da figura 4.16.

Sombras

Podemos observar então que os resultados até então são relativamente satisfatórios, conseguindo eliminar com sucesso grande parte das zonas de não estrada. Existem no entanto situações anormais que prejudicam o bom funcionamento deste sistema como é o caso da passagem em zonas de sombra.

Na figura 4.18 podemos verificar o impacto das zonas de sombra no processo de segmentação, que irá resultar posteriormente numa detecção errada da posição dos limites laterais. As figuras mostram sombras resultantes de objectos estáticos, neste caso um viaduto, que resultam numa mudança extrema nas características cromáticas da imagem, principalmente na componente de luminância.

O processo de detecção da faixa de rodagem não inclui nenhuma solução eficaz para eliminar correctamente as zonas de não estrada nestas situações, não sendo então possível fazer uma correcta detecção da zona de estrada. No entanto foi incluído neste processo um mecanismo que permite verificar quando um sensor de padrão se encontra em zona de sombra, por comparação com os restantes sensores. Este mecanismo permite então verificar quantos sensores estão posicionados neste tipo de zonas, passando esta informação para a tarefa seguinte: detecção dos limites laterais da faixa de rodagem. É importante realçar que este mecanismo tem apenas como objectivo identificar a passagem em zonas afectadas por sombra, passando esta mesma informação para a tarefa seguinte, informando do possível mau resultado do processo de detecção da faixa de rodagem.

A detecção das zonas sombreadas tem em conta o conhecimento adquirido e descrito na secção 4.3. Nessa secção é descrito que as zonas sombreadas são “mais azuis” que as restantes zonas, diferenciando-se também pelos seus valores mais baixos na componente de brilho, no espaço de cores HSV.

É exactamente este conhecimento que é utilizado para efectuar a detecção deste tipo de zonas. As características azuis numa região de análise podem ser detectadas após o cálculo da média da

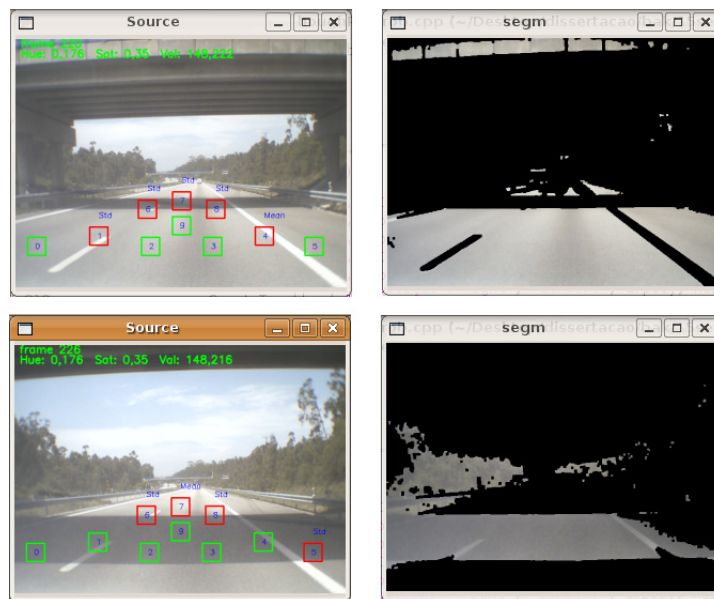


Figura 4.18: Impacto das sombras no processo de segmentação.

componente matiz e do seu respectivo desvio padrão. Uma região que revela características mais azuis do que as restantes irá certamente revelar um desvio padrão na componente matiz mais elevado. Isto acontece devido à representação do espaço de cor HSV na biblioteca usada. No OpenCV o azul é representado pelos valores mínimos da componente matiz, mas também pelos valores máximos, não esquecendo que o espaço de cor HSV é representado de forma circular. Ou seja, quando calculada a média de uma zona de análise na componente matiz, caso esta se situe em zona de sombra o resultado é um desvio padrão associado anormalmente alto.

A este procedimento junta-se também a comparação da variação na componente brilho, verificando se existem quedas acentuadas que possam significar que foram encontradas zonas de sombra.

Este processo pode então ser resumido ao seguinte:

1. Cálculo da média da componente matiz e do desvio padrão associado em cada sensor.
2. Cálculo da média dos desvios padrões de todos os sensores na componente matiz.
3. Cálculo da média da componente brilho em todos os sensores.

Estes valores calculados no passo 2 e 3 do processo em cima referido são utilizados para verificar em todas as frames se existe algum sensor que se caracterize por ter uma diferença acentuada nestas duas características: no desvio padrão da componente matiz e na média da componente brilho.

Por forma a melhor detectar estas variações não são apenas considerados os valores calculados em apenas uma frame mas sim nas últimas 20 frames. O que é feito então é calcular os valores descritos nos passos 2 e 3 do processo referido fazendo a sua média nas últimas 20 frames. De seguida, em cada frame são comparados os desvios padrões na componente matiz e a média da componente brilho de cada sensor, por forma a verificar se existem discrepâncias bruscas nestas duas características em algum dos sensores. Um sensor que se diferencie muito destas características simultaneamente, é considerado como estando situado em zona de sombra.

Desta forma é possível então passar para as tarefas posteriores a informação de que se detectou um sensor situado numa zona abrangida por sombra, que poderá assim funcionar como aviso sendo tratada esta informação da melhor forma possível.

Avaliação do processo

A figura 4.13 mostra a aplicação de 10 sensores padrão, em vez de apenas dois como foi inicialmente mostrado. Como é possível verificar, a segmentação destas imagens não elimina incorrectamente zonas da faixa de rodagem como acontecia anteriormente com apenas dois sensores. No entanto existem zonas da imagem irrelevantes, como por exemplo o céu (ver imagem superior direita da figura 4.13), que continuam a existir na imagem mesmo após a segmentação.

É por isso necessário ter atenção à quantidade de sensores, bem como ao seu tamanho em relação à imagem, uma vez que estas escolhas podem definir uma gama válida demasiado abrangente ou pelo contrário esta gama pode atingir um nível de especificidade que será difícil de corresponder e a segmentação irá ter maus resultados.

A localização destes sensores é também problemática uma vez que este trabalho foi realizado sem que existisse um setup de aquisição de imagem definido, sendo as experiências realizadas em sequências de vídeo que foram capturadas com dispositivos e setups muito diferentes.

Por este motivo fazer uma estimativa correcta da zona de estrada para posicionar correctamente os sensores padrão é uma tarefa complicada, uma vez que o campo de visão e as transformações efectuadas devido ao efeito perspectiva são muito diferentes em cada sequência de vídeo.

É possível no entanto mudar as posições destas zonas de análise em qualquer momento, estando já o sistema a funcionar, adaptando as suas localizações da melhor forma.

4.5 Detecção dos limites da faixa de rodagem

Apresentação do problema

Na secção anterior abordou-se a detecção da faixa de rodagem fazendo-se uma análise directa do seu padrão de cor. Nesta secção, faz-se uma abordagem diferente mas que resulta na mesma informação. Analisam-se mecanismos que permitem identificar os limites da faixa de rodagem. Em ambientes estruturados este limite é definido por uma linha delimitadora. Em cenários menos estruturados, os limites terão que ser identificados através da alteração de características da imagem entre a faixa de rodagem e o que estiver a partir do limite, como sejam rails, vegetação, etc.

Se os limites da faixa de rodagem forem eficientemente estimados, essa informação pode ser usada para localizar o veículo na estrada e mesmo para posicionar os sensores padrão abordados na secção anterior.

Abordagem inicial

Foram analisadas duas abordagens para realizar a detecção dos limites da faixa de rodagem. Na primeira abordagem utilizou-se a transformada linear de Hough. A transformada de Hough é um método para encontrar formas geométricas simples tais como linhas ou círculos numa imagem. Funciona apenas em imagens resultantes de processos de extracção de contornos; neste caso foi utilizado o algoritmo de Canny¹. No âmbito deste trabalho vamos apenas utilizar a transformada de Hough

¹As imagens do centro da figura 4.19 resultantes da aplicação do algoritmo de Canny estão representadas em negativo, para melhorar a visualização.

aplicada a linhas, uma vez que será esta a forma geométrica que é desejável encontrar.

A aplicação deste algoritmo resulta então num conjunto de linhas encontradas na imagem. No caso de uma imagem retirada de um ambiente real que mostra uma estrada, o resultado da transformada de Hough é um conjunto de linhas tangentes em cada ponto aos limites laterais da estrada (ver figura 4.19).

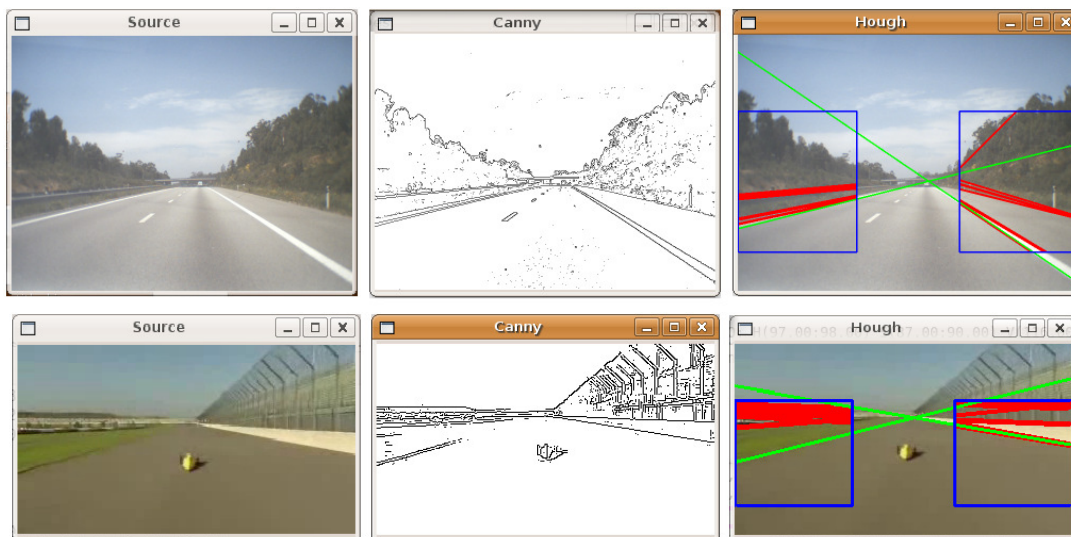


Figura 4.19: Aplicação da transformada de Hough: (esquerda) imagem original; (centro) imagem de contornos; (direita) imagem original com o resultado da aplicação da transformada de Hough.

Nas imagens do lado direito da figura 4.19 podemos verificar duas regiões delimitadas a azul que representam as zonas onde muito provavelmente se situam os limites laterais da estrada, sendo feita a análise apenas nestas regiões, uma vez que a aplicação da transformada de Hough em toda a imagem seria um desperdício de recursos. Podemos então verificar que no interior destas regiões existe um conjunto de linhas vermelhas, que representam todas as linhas detectadas pelo algoritmo de Hough.

Para conseguir obter um único resultado, ou seja, uma única linha que representasse da melhor forma a geometria da estrada, foi concebido um sistema de votação que em cada frame analisa todas as linhas resultantes, sendo capaz de ao fim de algumas frames estimar qual seria a melhor hipótese. Este sistema tem em consideração o número de vezes que cada linha aparece em cada posição, bem como a sua proximidade com a estrada, conseguindo desta forma identificar as linhas delimitadoras com razoável precisão. Nas imagens do lado direito da figura 4.19 o resultado deste mesmo processo está representado por uma linha verde, representando desta forma a posição mais provável da linha delimitadora da estrada.

Este modelo revela-se relativamente eficaz em estradas que não sejam constituídas por muitas curvas. O problema fundamental é que o método de Hough foi desenvolvido para detectar apenas rectas, o que mais tarde se tornou um problema quando foi necessário detectar e compreender as situações de curvatura da estrada.

A transformada de Hough revelou-se também um processo relativamente pesado computacionalmente, mesmo após a definição de zonas de interesse para evitar a sua aplicação à imagem completa. A sua aplicação nas sequências de vídeo atrasava claramente o processamento de cada imagem, podendo este ser também apontado como um ponto negativo na sua utilização num sistema de aquisição de imagem em tempo real.

Após esta primeira abordagem, optou-se por uma nova metodologia que pudesse permitir a detecção dos limites da estrada transmitindo informação de curvatura.

Na segunda abordagem manteve-se o uso da imagem de contornos, mas substituiu-se o uso da transformada linear de Hough por um método mais leve computacionalmente. Assumindo que se conhece uma vizinhança do limite da faixa, usam-se sensores de linha para detectar esse limite. Sensores de linha são linhas traçadas sobre uma imagem e segundo as quais a imagem é analisada à procura de alguma particularidade. Neste caso a particularidade procurada é o início do contorno.

A figura 4.20 ilustra o uso deste tipo de sensores. Um pente de linhas sobre a imagem, numa posição estimadamente paralela ao contorno que se pretende detectar é usado para, quando varrido da direita/baixo para a esquerda/cima, detectar o início do contorno. Tal como é possível verificar com mais detalhe na figura, cada linha do sensor detecta a primeira transição de cor, de branco para preto², ignorando a segunda linha que se encontra nesta imagem e que está também na zona do sensor.

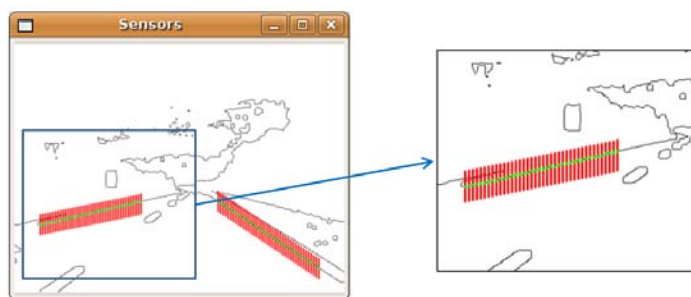


Figura 4.20: Sensores de linha.

Ajuste dinâmico dos sensores

Os sensores de linha foram inicialmente criados para que a sua disposição fosse apenas horizontal ou vertical. No entanto, devido à natureza deste problema, foram introduzidos mecanismos que possibilitaram a alteração da disposição horizontal ou vertical para uma disposição diagonal que possibilitasse o ajuste da sua inclinação conforme o desejado. Desta forma é possível regular a orientação dos sensores de acordo com a inclinação dos limites da faixa de rodagem em cada momento.

Tal como já foi dito, as faixas de rodagem onde o sistema se destina a trabalhar não têm largura fixa. Por este motivo é necessário encontrar uma solução que permita de forma automática procurar a melhor posição para cada sensor, estimando assim a localização provável dos limites da faixa de rodagem. Para estabelecer a posição inicial dos sensores é feita uma pesquisa para tentar perceber onde serão os limites laterais da faixa de rodagem. Essa pesquisa consiste num varrimento horizontal e vertical, tal como se mostra na figura 4.21, de forma a tentar encontrar a sua posição mais provável. Este processamento consiste na criação de um sensor temporário que efectua este varrimento na ordem indicada, para o sensor esquerdo e para o direito, até ser possível detectar alguns pontos que indiquem onde provavelmente se situa o limite da faixa de rodagem.

Logo de seguida os sensores são posicionados nesta mesma localização. Este mesmo processo é repetido sempre que algum dos sensores deixe de conseguir detectar correctamente a posição da faixa

²A imagem da figura 4.20 é a imagem em negativo da original, para melhorar a visualização. Desta forma quando se diz que os sensores detectam a primeira transição de branco para preto, na verdade a detecção é feita na primeira transição de preto para branco.

de rodagem, o que é possível acontecer com uma mudança brusca de ambiente, ou devido a uma zona de sombra.



Figura 4.21: Pesquisa da posição óptima dos sensores.

Uma vez que a posição do veículo em relação à estrada vai mudando em função do tempo, é impossível definir uma posição e orientação estática dos sensores por forma a que estes estejam sempre correctamente posicionados para detectar os limites laterais. Desta forma foi introduzido um mecanismo que possibilita a adaptação dinâmica da posição e orientação destes mesmos sensores à medida que o veículo se vai deslocando.

Após ter sido encontrada a posição inicial de cada sensor é necessário que estes se ajustem dinamicamente em posição e orientação para melhor conseguir detectar os limites da faixa de rodagem. Este mecanismo baseia-se nos pontos que são encontrados pelos sensores, tentando sempre que estes pontos se encontrem exactamente a meio de cada linha do sensor. Ou seja, este processo consiste em verificar a posição dos pontos encontrados relativamente à posição do sensor, tentando em cada momento posicionar os sensores por forma a que os pontos encontrados estejam numa posição intermédia de cada sensor.

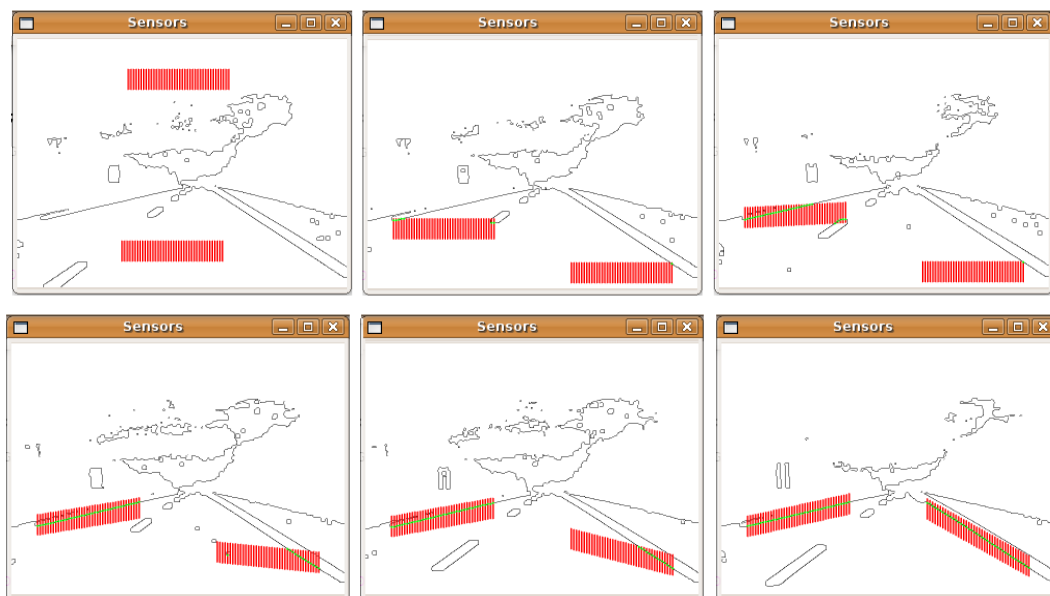


Figura 4.22: Ajuste dinâmico da posição dos sensores de linha. Da esquerda para a direita, de cima para baixo: frames 23,24,27,30,36 e 40 de uma mesma sequência de vídeo.

Como exemplo podemos verificar as imagens na figura 4.22. Estas imagens fazem parte da mesma sequência de vídeo, espaçadas por 17 frames³. Na frame inicial, os sensores são propositadamente postos manualmente em sítios onde nenhuma faixa irá ser encontrada para testar a sua capacidade de auto-localização. Logo na frame seguinte os dois sensores encontram uma posição onde já se encontra parte dos limites da faixa de rodagem. O sensor da direita consegue-se auto-posicionar em cerca de apenas 7 frames enquanto que o sensor da esquerda leva um pouco mais, cerca de 17 frames, até estar completamente ajustado à faixa lateral.

Neste processo é possível então verificar que inicialmente é feito o varrimento descrito na figura 4.21, após o qual os sensores são localizados na melhor posição. De seguida existe um mecanismo que auto-ajusta a sua posição e inclinação de acordo com as faixas delimitadoras, sem qualquer intervenção humana.

Estimativa dos limites laterais da faixa

No fim desta fase tem-se então como resultado um conjunto de pontos, que representam os limites laterais da estrada. É necessário de alguma forma agrupar ou ligar estes pontos, por forma a ser possível compreender o conjunto como um todo.

O ideal seria encontrar um processo que dado um conjunto de pontos pudesse de certa forma criar um único resultado, que ponderasse as posições de todos os pontos, tendendo a eliminar algumas possíveis variações incorrectas. A solução encontrada foi a utilização de uma regressão linear. Este processo determina então uma solução que tem em conta a posição de todos os pontos, tendendo a eliminar possíveis pontos que tenham sido incorrectamente detectados

Dado um conjunto de pontos (x, y) definidos no espaço, através de uma regressão linear é possível obter uma equação do tipo $y = mx + b$ que resultará numa recta que deve definir uma boa aproximação ao conjunto de pontos (ver figura 4.23). A regressão linear é feita por forma a que esta mesma linha seja tão próxima dos pontos quanto possível, ou seja, deve resultar numa situação que minimize os resíduos. Os resíduos são definidos pelas diferenças no eixo Y entre cada um dos pontos iniciais e a recta definida pela regressão linear.

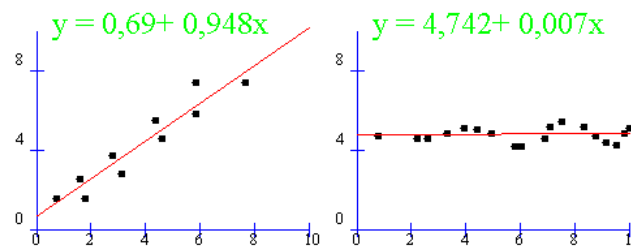


Figura 4.23: Exemplo de regressões lineares.

A forma mais comum de regressão linear, que será utilizada neste caso específico, é baseada no método dos mínimos quadrados. O nome é relacionado com o critério para encontrar a melhor solução, que é minimizar a soma do quadrado dos residuais. Ou seja, a regressão dos mínimos quadrados é representada por uma linha que minimiza a soma dos quadrados dos residuais de cada ponto [63].

No caso deste trabalho não fazia sentido fazer uma única regressão linear ao conjunto global de pontos detectados, uma vez que iria resultar em apenas uma recta o que iria ser um mau resultado uma

³As imagens da figura 4.22 são as imagens em negativo das originais, uma vez que desta forma se melhora a visualização.

vez que queremos um resultado com alguma definição, que possa por exemplo representar situações de curvatura.

Por este motivo para cada ponto é calculada uma regressão linear dos seus 9 pontos mais próximos, até chegar a um conjunto que inclua o último ponto. Ou seja, num conjunto de 15 pontos encontrados, o resultado serão 6 rectas resultantes da aplicação do método de regressão linear aos intervalos entre os pontos: [1-10],[2-11],[3-12],[4-13],[5-14] e [6-15].

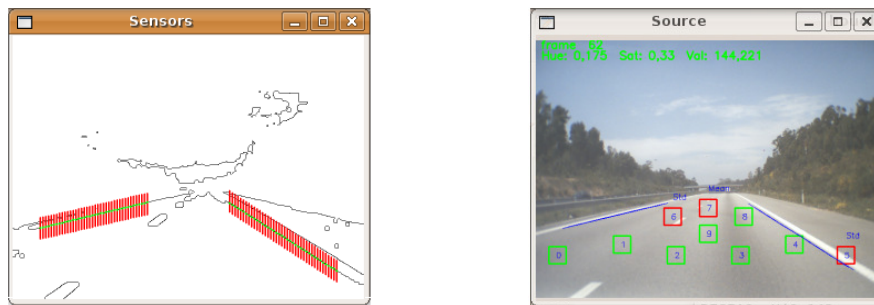


Figura 4.24: Sensores de linha e regressão linear dos pontos encontrados.

Na imagem do lado direito da figura 4.24 estão várias rectas representadas a azul que resultam de várias regressões lineares dos pontos encontrados pelos sensores de linha. Neste caso, como se trata de um conjunto de pontos que estão correctamente alinhados, o resultado é um conjunto de rectas que se assemelham a uma só. No caso da figura 4.25 podemos já verificar que o resultado das regressões lineares representam uma ligeira curvatura, uma vez que é essa a tendência dos pontos detectados pelos sensores de linha.

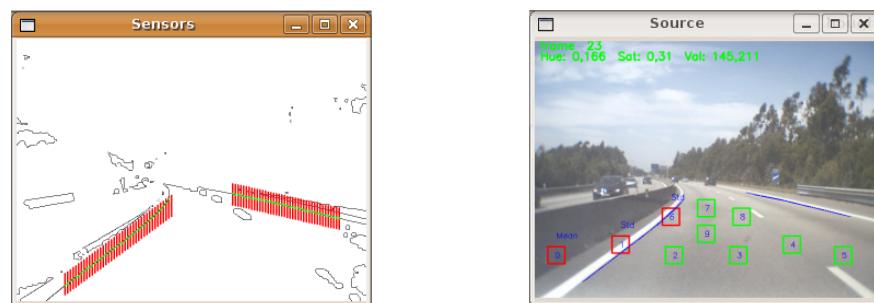


Figura 4.25: Sensores de linha e regressão linear dos pontos encontrados.

Desta forma é possível observar este conjunto de rectas resultantes das regressões lineares como uma estimativa dos limites laterais da faixa de rodagem, quer seja em situações de curvatura ou não.

Impacto das sombras

Tal como já foi referido na secção 4.4, a passagem em zonas fortemente atingidas por sombra causa um resultado incorrecto na tarefa de detecção da faixa de rodagem. A imagem resultante deste processo nestas situações é portanto inútil para a correcta detecção dos limites laterais. No entanto, um mecanismo construído no processo de detecção da faixa de rodagem, descrito na secção 4.4, permite que seja feita a detecção de zonas afectadas por sombra, sendo possível saber quando existem sensores padrão que estão situados em zonas de sombra. Quando dois ou mais sensores do processo

de detecção da faixa de rodagem se encontram em zonas de sombra, é considerado então que se está a atravessar uma zona com forte impacto de sombra.

A correcta detecção deste tipo de zonas permite ao sistema de detecção dos limites laterais ter a percepção de quando se encontra numa zona com forte impacto de sombra, por forma a alterar o seu método de actuação normal. A solução encontrada para minimizar este problema foi manter a última estimativa correcta dos limites laterais da faixa (ver figura 4.26), enquanto uma zona de sombra é atravessada.

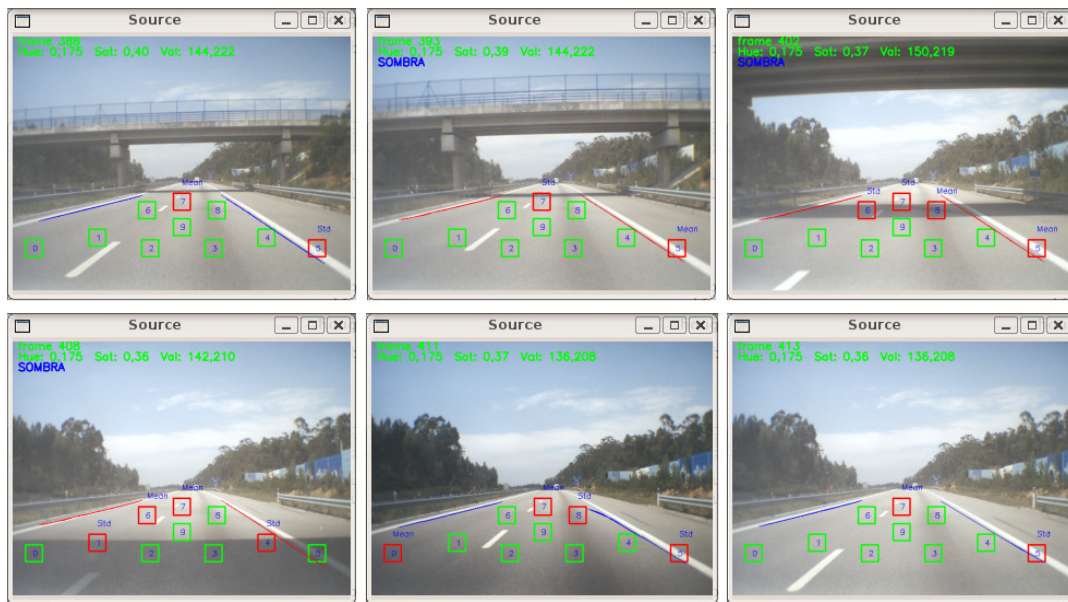


Figura 4.26: Imunidade do processo de detecção dos limites laterais relativamente às sombras. Da direita para a esquerda, de cima para baixo: frames 388, 393, 402, 408, 411 e 413 da mesma sequência de vídeo.

Podemos verificar que a frame 388 na figura 4.26 representa um detecção correcta dos limites laterais, detecção esta que é mantida até à frame 411 como estimativa da sua possível posição, uma vez que nas frames intermédias se encontra uma zona sombreada e não é portanto possível a sua correcta detecção.

A cor dos limites da faixa de rodagem indica isto mesmo, mudando de azul para vermelho quando se trata de uma zona sombreada em que os limites laterais não são detectados mas sim estimados da última frame sem zonas sombreadas, onde foi feita a sua correcta detecção.

Assim, não sendo possível efectuar a correcta detecção dos limites laterais em zonas com sombra, é feita uma estimativa destes mesmos limites que se espera ser aproximada da realidade. Esta estimativa apresenta-se normalmente eficaz, excepto em casos de mudança brusca da posição dos limites laterais da faixa, enquanto é atravessada uma zona com sombra.

4.6 Detecção de obstáculos

Apresentação do problema

A detecção de determinados objectos em zonas aleatórias da imagem e o seu seguimento é por si só uma tarefa complexa na área da visão computadorizada. A heterogeneidade de características

apresentadas pelos veículos comuns são uma dificuldade acrescida deste processo, uma vez que são estes os objectos que será necessário detectar.

Existem ainda várias condicionantes num ambiente real que dificultam esta tarefa. Por exemplo, os veículos podem variar em forma, tamanho e cor, a aparência do veículo depende da sua orientação e localização sendo afectada por objectos próximos que podem esconder partes deste mesmo veículo. A iluminação é sempre um factor a ter em conta, uma vez que é imprevisível, bem como as interacções entre os vários veículos que circulem na mesma faixa. O ruído muitas vezes presente no plano de fundo ou nas partes laterais da imagem complicam também a tarefa de detecção de veículos.

Abordagens comuns ao problema

Devido a estas condicionantes é necessário que existam mecanismos que possam numa primeira fase encontrar na imagem possíveis veículos, sendo possível de alguma forma validar nas imagens seguintes se se trata realmente de um veículo ou de uma detecção errada. É por este motivo que grande parte da literatura descreve métodos que seguem dois passos essenciais: (1) geração de hipóteses e (2) verificação destas mesmas hipóteses [54, 37].

O primeiro passo consiste numa geração de várias hipóteses de localizações de possíveis veículos. Este primeiro passo consiste no aproveitamento de uma série de heurísticas conhecidas dos veículos, que permitem formular estas mesmas hipóteses.

O segundo passo consiste na verificação ao longo de várias frames da veracidade destas mesmas hipóteses, efectuando vários testes para eliminar as hipóteses erradas e consequentemente eleger as correctas (ver figura 4.27).

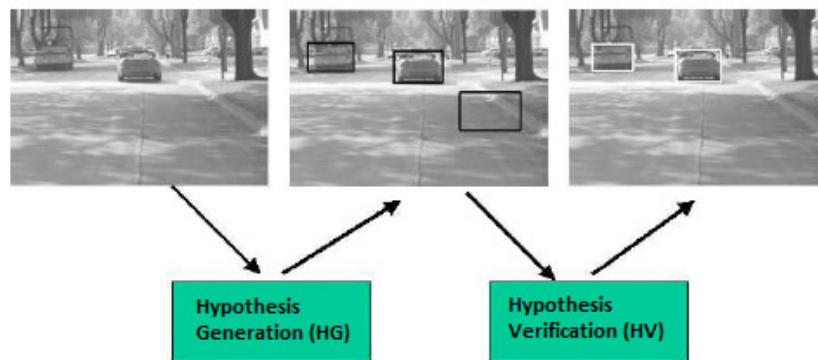


Figura 4.27: Exemplo do método de geração e verificação de hipóteses [37].

Para elaborar um sistema de detecção de veículos é necessário então que existam várias características pré-conhecidas que permitam fazer a validação das hipóteses encontradas.

Os vários sistemas de detecção de obstáculos que são propostos na generalidade da literatura podem ser classificados em três grandes categorias: 1) baseados no conhecimento, 2) baseados em imagens estereoscópicas e 3) baseados no movimento.

Os sistemas baseados em imagens estereoscópicas tiram partido de um sistema de aquisição com duas câmaras que permite posteriormente ter uma noção de 3D, para identificar possíveis obstáculos. Estes sistemas têm uma necessidade computacional normalmente elevada, bem como revelam por vezes problemas com as vibrações das câmaras para depois fazer a correspondência entre as duas imagens adquiridas.

Os sistemas baseados em movimento tiram partido da dinâmica dos veículos para fazer a sua detecção, capturando assim as diferenças entre imagens consecutivas para atingir o objectivo. Um vector é gerado para cada ponto na imagem para medir as diferenças de posições entre cada frame.

Os sistemas baseados no conhecimento apoiam-se nas várias características comuns na generalidade dos veículos: simetria, cor, sombras, cantos, contornos verticais/horizontais, texturas, luzes dos veículos, etc. Uma vez que normalmente estas características se encontram presentes nos veículos como em nenhum outro lado, é possível a sua utilização para distinguir os veículos de outros objectos.

Deteção de veículos baseada em conhecimento

Este trabalho destina-se a um sistema com apenas uma câmara, sendo por isso impossível a utilização de metodologias que utilizam imagens estereoscópicas. Os sistemas baseados em movimento são computacionalmente bastante pesados, bem como revelam muito pouca sensibilidade na detecção a baixas velocidades, sendo mais utilizados em sistemas com velocidades relativas mais altas.

Será portanto usada uma abordagem a este problema baseada em métodos que utilizam o conhecimento para detecção de veículos. Este tipo de sistemas aproveitam características básicas comuns a todos os veículos, que normalmente somos capazes de identificar, para fazer a sua detecção.

Existem várias características que definem um veículo e estas são procuradas sequencialmente nas imagens para validar a detecção de um veículo. O maior defeito de um sistema que utilize esta metodologia é que se as características não forem suficientemente visíveis, devido a ruído ou oclusão parcial de outros veículos, o veículo pode não ser encontrado [37].

As características normalmente mais usadas nestes tipos de sistemas são as seguintes:

1. Contornos verticais/horizontais
2. Simetria
3. Cor e textura
4. Sombra
5. Luzes dos veículos
6. Forma rectangular, cantos

De seguida irá ser feito um estudo sobre as características anteriormente referidas, do ponto de vista prático, por forma a verificar quais serão as que melhores resultados podem fornecer para fazer a detecção de veículos. Para este efeito serão desenvolvidos alguns algoritmos por forma a possibilitar a extracção destas mesmas características, usando sempre que possível as potencialidades do OpenCV.

Após o estudo será proposta uma solução que, tendo em conta os resultados do estudo, utiliza as características comuns nos veículos para fazer a sua detecção.

Estudo das características comuns nos veículos para utilização em processos de detecção

Contornos verticais/horizontais

Muitas abordagens para interpretação de imagens são baseadas em algoritmos de extracção de contornos, uma vez que estes têm como objectivo realçar as características mais importantes da imagem simplificando a sua análise posterior. A detecção de contornos realça o contraste da imagem, as

mudanças de intensidade, resultando assim numa imagem que realça as características estruturais dos objectos.

Diferentes pontos de vista de um veículo, especialmente de frente ou de trás, contêm bastantes estruturas horizontais e verticais, tais como as janelas, o pára-choques, a matrícula e a sombra por baixo do veículo. Por este motivo a extracção de contornos numa imagem com objectivo de detecção de veículos torna-se uma tarefa normalmente necessária. Na figura 4.28 pode-se observar que tanto o veículo da esquerda como o da direita se caracterizam por ter traços verticais e horizontais bastante fortes.

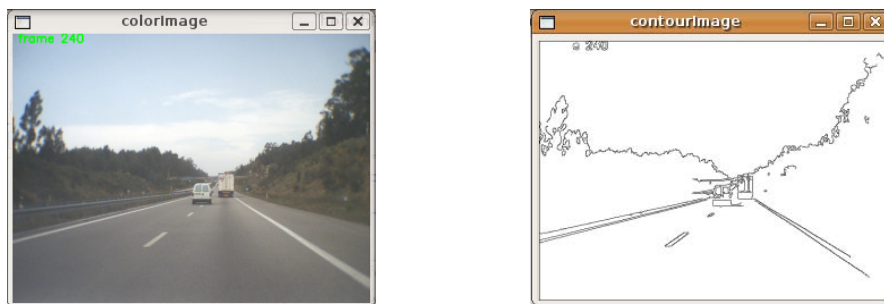


Figura 4.28: Aplicação de um algoritmo de extracção de contornos, Canny.

Desta forma e aproveitando os métodos de detecção de contornos do OpenCV várias experiências foram feitas de forma tentar detectar características na imagem que determinassem a presença de um veículo.

A aplicação do algoritmo de Canny resulta numa extracção de todos os contornos que se encontram relativamente bem definidos (ver imagem do lado direito da figura 4.28). Nas duas imagens da figura 4.29 podemos observar que a aplicação do algoritmo de Sobel resulta na extracção das características horizontais e verticais separadamente. Ambos os algoritmos, Canny e Sobel, são aplicados na primeira imagem da figura 4.28 transformada em tons de cinzento, ou seja, monocromática.

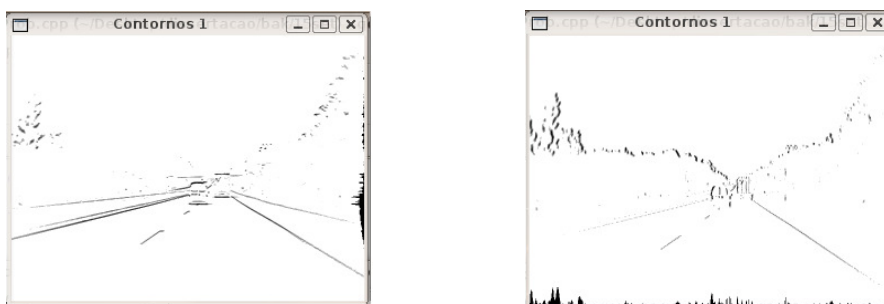


Figura 4.29: Aplicação do algoritmo de Sobel para extracção horizontal e vertical de contornos à imagem

Para melhor perceber onde seriam as características verticais e horizontais dominantes na imagem foi adicionado um histograma a cada uma das componentes, representado nas próprias imagens de extracção. Na extracção horizontal pode-se verificar que os picos do histograma apontam claramente para a posição do veículo. Na extracção vertical existe bastante ruído na parte esquerda da imagem, devido à vegetação circundante, o que faz com que o pico do histograma se situe nesta zona da imagem, não contribuindo assim para a localização eficaz do veículo. No entanto se a imagem fosse alvo de algum processamento anterior para retirar este tipo de objectos que não pertencem à faixa

de rotação, os picos do histograma estariam certamente posicionados na zona onde se encontra o veículo.

A aplicação deste algoritmo tem resultados bastante interessantes principalmente na extracção dos contornos horizontais, onde é possível verificar que os contornos mais bem definidos são os do próprio veículo.

Além deste processamento feito pelo algoritmo de Sobel existem ainda outras formas de identificar onde se encontram as características horizontais e verticais dominantes.

Uma forma alternativa de o fazer foi através do cálculo de um gradiente vertical e horizontal dos valores de cada píxel da imagem. O cálculo destes gradientes é feito utilizando o resultado da aplicação do algoritmo de Canny.

Desta forma para calcular o gradiente vertical da imagem o processo resume-se a determinar para cada píxel da imagem resultante, uma intensidade que resulta da média das intensidades de todos os píxeis que são verticalmente vizinhos, dentro de um certo intervalo, deste mesmo píxel na imagem fonte (ver figura 4.30).

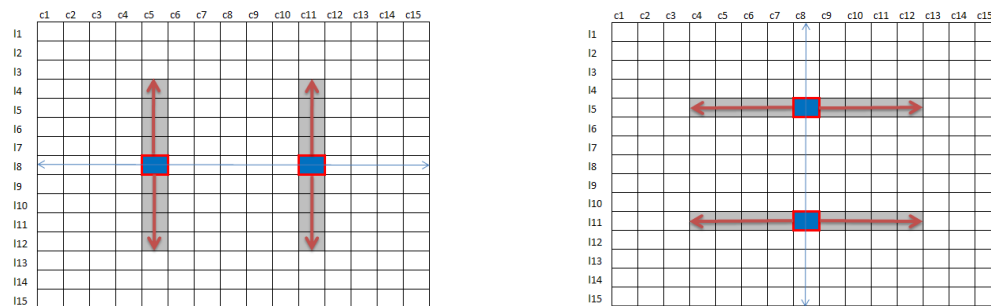


Figura 4.30: Cálculo dos gradientes vertical e horizontal

Na imagem do lado esquerdo da figura 4.30 mostra-se o cálculo do gradiente vertical de apenas dois dos píxeis da imagem resultante. Neste caso o píxel situado na linha 8 e coluna 5 da imagem resultante, irá ser determinado a partir da média das intensidades dos 8 píxeis verticalmente mais próximos, na imagem fonte. O intervalo de 8 píxeis foi experimentalmente encontrado como uma das melhores soluções.

A integração horizontal é um processo análogo à integração vertical, sendo que a única diferença é que a imagem resultante é determinada a partir da média dos píxeis horizontalmente vizinhos dos píxeis da imagem fonte.

Nas imagens da figura 4.31 podemos visualizar o resultado do cálculo dos gradientes horizontal e vertical, bem como os histogramas correspondentes. No histograma na imagem da esquerda podemos verificar que horizontalmente os resultados não são muito expressivos, mas no entanto os picos do histograma apontam correctamente para a localização do veículo se verificarmos a imagem mais atentamente. Quanto ao gradiente vertical, na imagem da direita, os resultados são bastante expressivos uma vez que os picos do histograma apontam claramente para a posição dos veículos.

Se compararmos os dois tipos de processamentos para extracção de contornos horizontais e verticais, representados pelas figuras 4.29 e 4.31, é possível verificar que ambos contribuem positivamente para obter bons resultados na localização dos veículos, excepto a última imagem da figura 4.29 que erradamente tem o pico do seu histograma numa zona que nada tem a ver com a localização dos veículos.

Desta forma é possível dizer que este tipo de processamento é fundamental para este tipo de abordagem na detecção de veículos, uma vez que os contornos horizontais e verticais são de facto uma

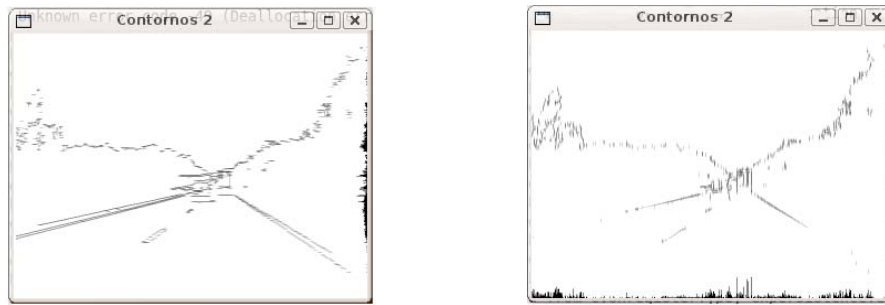


Figura 4.31: Cálculo dos gradientes vertical e horizontal da figura 4.28

característica presente na generalidade dos veículos. É no entanto necessário algum processamento anterior para tentar remover da imagem certos objectos tais como a vegetação ou sinais na estrada, que, se forem caracterizadas por algum tipo de traços verticais ou horizontais, podem contribuir para a errada detecção utilizando estes processos de extracção.

Simetria

A simetria é uma das características presentes em quase todos os objectos feitos pelo homem. Os veículos são disto um exemplo, revelando bastantes traços simétricos, principalmente quando vistos pela frente ou por trás.

Para analisar as características simétricas de uma imagem, a análise da imagem em tons de cinzento não é suficiente. As reflexões podem causar eixos de simetria errados, bem como as áreas homogéneas e os padrões de fundo. Como tal, o mais eficiente é fazer um primeiro processamento para extrair os contornos verticais e horizontais por forma a eliminar a presença de áreas homogéneas e da generalidade das reflexões [64].

Dando continuidade ao trabalho anterior, o algoritmo elaborado para detecção de simetria irá ser aplicado aos resultados dos processos de extracção de contornos mostrados nas figuras 4.29 e 4.31, resultantes da aplicação do algoritmo de Sobel e do cálculo dos gradientes horizontal e vertical.

O algoritmo de detecção de simetria horizontal resume-se ao varrimento vertical da imagem, por forma a encontrar pares de píxeis na mesma coluna em que a distância entre eles seja menor que um intervalo pré-definido. Quando pares de píxeis são encontrados, com valor de intensidade relevante, um píxel no meio destes dois píxeis é marcado na imagem resultante. Aplicando este processo a toda a imagem, o resultado é uma relevância de todas as características de simetria horizontal da imagem (ver primeira imagem das figuras 4.32 e 4.33).

O processo de detecção de simetria vertical é análogo a este mesmo processo (ver segunda imagem das figuras 4.32 e 4.33).

A figura 4.32 resulta da aplicação do processamento anteriormente descrito às imagens da figura 4.29. Tal como é possível verificar, o resultado tanto horizontal como verticalmente não é muito expressivo. Horizontalmente existe alguma tendência de simetria na localização original dos veículos, principalmente do veículo do lado esquerdo, não se notando no entanto nenhum eixo de simetria bem definido. Verticalmente as características simétricas encontradas estão algo dispersas não sendo visível nenhum eixo de simetria bem definido. No entanto se fosse aplicado algum processamento prévio para retirar da imagem os objectos que se localizam fora da faixa de rodagem, os resultados de simetria vertical iriam ser apenas os do veículo, uma vez que as restantes características simétricas encontradas são devido à vegetação circundante.

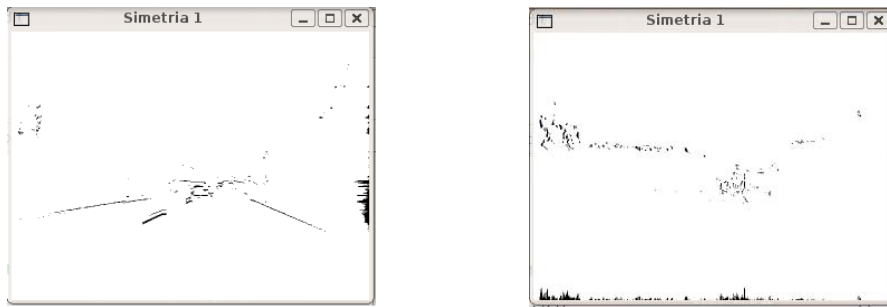


Figura 4.32: Cálculo da simetria das imagem na figura 4.29.

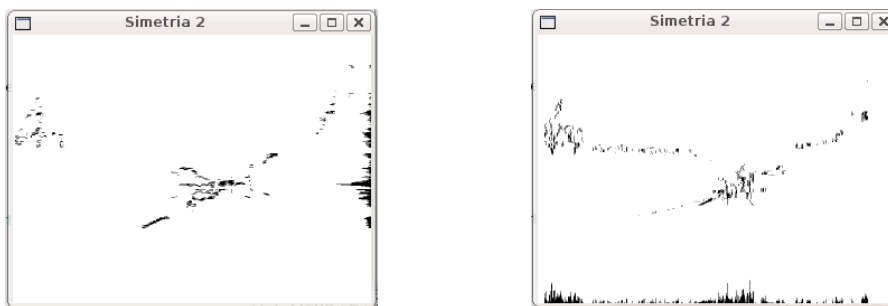


Figura 4.33: Cálculo da simetria das imagens na figura 4.31.

A figura 4.33 é resultado da extracção de simetrias verticais e horizontais das imagens da figura 4.31. Na detecção horizontal é possível verificar que existem dois eixos de simetria correspondentes aos dois veículos. Os traços horizontalmente simétricos encontrados, ainda que contenham algum ruído, possibilitam assim estimativa da localização dos veículos. Verticalmente os traços simétricos encontrados são razoavelmente visíveis. Existem três picos no histograma correspondente, que assinalam então a simetria vertical de cada veículo e ainda um terceiro traço simétrico detectado que corresponde à vegetação lateral.

A simetria como forma de detecção de veículos é uma mais valia, tal como já tinha sido referido, e é utilizada em vários sistemas presentes na literatura. Os resultados apresentados nesta secção são prova disso mesmo, conseguindo através da extracção de simetria fazer uma estimativa da posição de possíveis obstáculos. No entanto a simetria é normalmente usada como forma de complementar outras formas de análise, uma vez que é uma característica presente em muitos objectos sendo natural que muitas vezes a detecção de fortes características simétricas provenham de outros objectos que não os próprios veículos.

Cor e textura

A cor é uma característica importante num veículo sendo o que normalmente mais o distingue do resto dos objectos numa imagem. No entanto esta propriedade é também de difícil utilização uma vez que os veículos existentes apresentam uma gama muito variada de cores, tornando-se assim complexo utilizar esta característica para a sua detecção. Além disto a cor depende sempre das características do sensor, da iluminação e brilho de um objecto, fazendo com que a mesma cor tenha características diferentes de acordo com a altura do dia ou mesmo com a orientação do veículo.

Uma abordagem possível é fazer uma segmentação da estrada para tentar isolar tudo o que nela se situe com características diferentes, tentando localizar possíveis veículos. Não existem no entanto

muitos sistemas descritos na literatura que se baseiam na cor para fazer esta mesma detecção.

A textura é um conceito bastante abstracto, normalmente atribuída à percepção humana da cor de um certo objecto. Quando se olha para uma imagem é fácil agrupar regiões que tenham as mesmas propriedades de cor em objectos. Não existe nenhuma definição matemática para definir o conceito de textura. No entanto, a repetição de um padrão ou de padrões numa região é certamente uma característica de uma textura. A variabilidade de forma, tamanho, cor e orientação destes mesmos padrões dificulta a sua localização principalmente quando se trata de visão computadorizada [65, 66].

Sombra

A sombra é uma característica presente em todos os veículos num ambiente real, caso se trate de uma situação com luminosidade regular e sem intervenção de sombras de outros objectos, sendo uma característica bastante importante para formular uma hipótese da localização de um possível veículo.

Baseado no facto de a área por baixo de um veículo ser substancialmente mais escura que qualquer outra parte da estrada, é possível olhar para a imagem de baixo para cima esperando encontrar transições de zonas mais brilhantes para zonas menos brilhantes, para encontrar possíveis sombras. É uma característica que depende directamente da luminosidade da imagem, ou seja, das condições climáticas, sendo desta forma difícil de definir um padrão que defina correctamente as sombras.

Para segmentar zonas de sombra é assim necessário ter em atenção as características da imagem em cada momento, definindo dinamicamente os thresholds baixo e alto que permitam assim isolar as sombras.

Utilizando o conhecimento adquirido na secção 4.3, foram escolhidos dois métodos de detecção de sombras: detectar diminuições bruscas da componente brilho do espaço de cores HSV e detectar diminuições bruscas nas três componentes, simultaneamente, no espaço de cores RGB. Para detectar estas transições a solução encontrada foi fazer um varrimento vertical da imagem, de baixo para cima ao longo de todas as suas colunas, por forma a encontrar estas transições bruscas, tanto num como noutro método.

Definir um intervalo estático que definisse se a transição era suficientemente grande para classificar a zona como sombra não é o mais correcto uma vez que os ambientes podem variar muito, e é difícil definir um valor estático que funcione para todas as situações. A solução encontrada foi calcular em cada imagem a média das componentes que interessam, dependendo do método usado (brilho no HSV e as três componentes no RGB), à medida que é feito o varrimento vertical e ir verificando se o píxel seguinte contém um valor que seja muito mais baixo do que a média calculada até então.

A relação encontrada para o fazer é verificar se o píxel seguinte tem uma queda na(s) componente(s) que seja superior a $\frac{1}{4}$ da média calculada até então. O resultado deste processamento é apresentado pelas imagens do centro e da direita da figura 4.34.

Existem também algumas zonas da vegetação circundante que são erradamente detectadas como sombra, mas na faixa de rodagem a única zona que é detectada é a zona por baixo do veículo precedente que diz respeito à sua sombra (ver figura 4.34).

A sombra é uma característica bastante usada nas metodologias de detecção de veículos descritas na literatura, uma vez que devido às suas características particulares é possível fazer a sua distinção com algum sucesso.

A detecção de sombras é muitas vezes utilizada como método inicial de detecção, para formular as hipóteses de possíveis localizações de veículos, sendo essas hipóteses posteriormente avaliadas recorrendo a outro tipo de características.

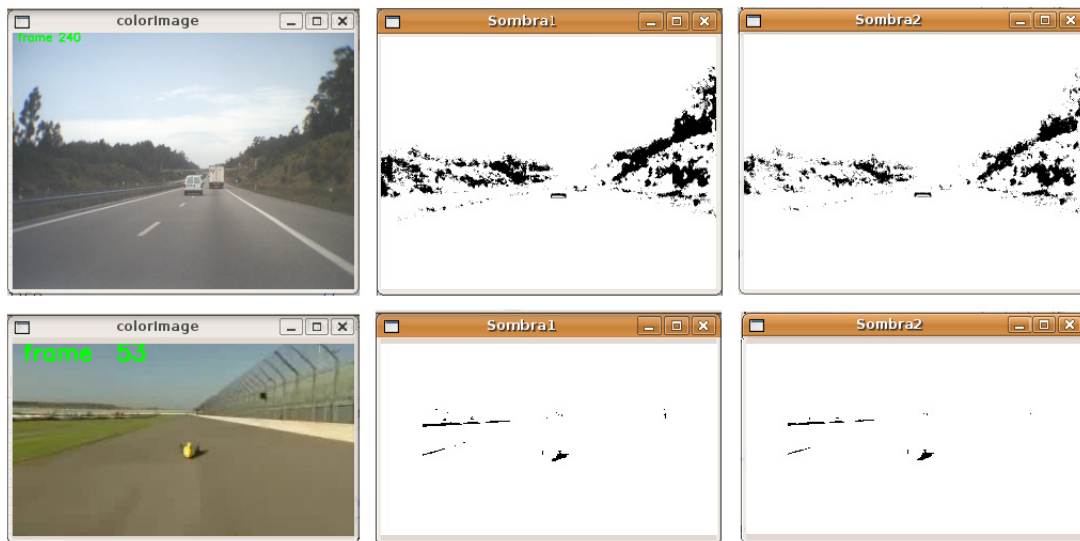


Figura 4.34: Detecção de sombras utilizando a componente brilho(HSV) e a componente azul(RGB) respectivamente

Luzes dos veículos

As luzes são características presentes em todos os veículos, apresentando normalmente uma cor e localização semelhantes. De todas as características descritas anteriormente apenas esta é utilizável em ambientes nocturnos, uma vez que num ambiente com pouca ou nenhuma luminosidade, a detecção das outras características torna-se muito difícil ou quase impossível.

Em ambientes em que exista uma boa iluminação a detecção das luzes de um veículo não é muito comum uma vez que já não apresentam uma cor tão distinguível, existindo também o problema de a cor do veículo se aproximar da cor das próprias luzes. No entanto foi feita a tentativa de detectar a presença das luzes de um veículo, num ambiente regularmente iluminado, recorrendo à cor que normalmente as caracterizam, o vermelho. Na figura 4.35 podemos ver um exemplo razoavelmente bem sucedido da tentativa de segmentar as luzes do veículo. As luzes traseiras do veículo ficam realmente bastante visíveis quando se elimina todas as outras regiões que não são compostas por vermelho.

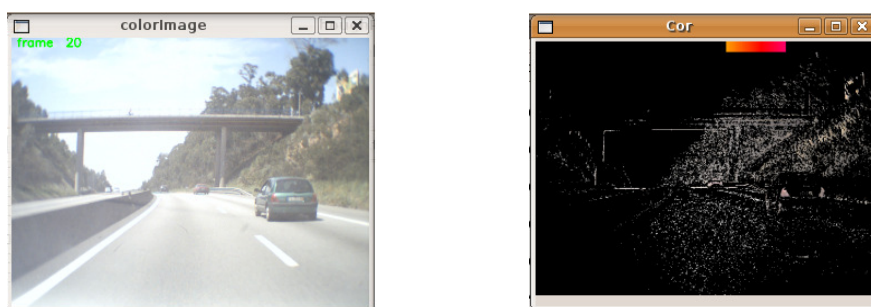


Figura 4.35: Segmentação de cor para encontrar as luzes dos veículos

O problema é que a estrada, uma vez que a sua cor dominante é o cinzento, contém uma variação muito grande na componente matiz, tal como já tinha sido referido anteriormente. Por este mesmo motivo a tentativa de segmentação de qualquer cor específica neste espaço de cores normalmente

resulta sempre na detecção de algumas zonas de estrada, tal como acontece na figura 4.35 e 4.36 quando se tenta segmentar o vermelho.

Na figura 4.36 o que acontece é que o veículo é vermelho, ou seja, da mesma cor das próprias luzes. O que acontece neste caso, tal como seria de esperar, é que a zona segmentada abrange todo o veículo.



Figura 4.36: Segmentação de cor para encontrar as luzes dos veículos

A detecção das luzes de um veículo não é certamente dos métodos mais usados para a detecção de veículos, podendo no entanto ser mais um elemento para avaliação. A sua cor e posição mudam de veículo para veículo não sendo assim intuitivo fazer a sua detecção.

Forma rectangular, cantos

Os veículos normalmente apresentam uma forma rectangular, principalmente quando vistos de trás, sendo também possível verificar a existência de quatro cantos. Existem alguns métodos descritos na literatura que juntamente com outro tipo de informação usam estas características de forma complementar no processo de detecção.

Na figura 4.37 podemos verificar o processo de extracção de cantos. Este processo utiliza o resultado da detecção de contornos verticais e horizontais, anteriormente mencionada, extraindo as intersecções entre estes dois resultados (contornos verticais e horizontais), do qual resultam vários pontos. Ambas as imagens na figura 4.37 foram alvo de um processo de dilatação para tornar os resultados mais visíveis.



Figura 4.37: Extracção de cantos utilizando a intersecção dos contornos verticais e horizontais, calculados com o algoritmo de Sobel (imagem da esquerda) e cálculo dos gradientes (imagem da direita).

A imagem da esquerda da figura 4.37 é resultado da intersecção entre os contornos verticais e horizontais usando o algoritmo de Sobel, enquanto que a segunda imagem resulta do mesmo processo mas os contornos são calculados através dos gradientes. Podemos verificar que na primeira imagem

da figura 4.37 existe uma nuvem de pontos não muito bem definida na zona onde se encontram os veículos, existindo ainda vários pontos que resultam dos limites laterais da faixa e da vegetação. Na segunda imagem podemos verificar que os pontos encontrados convergem para a localização dos veículos. Vários pontos são detectados uma vez que a estrutura dos veículos possui várias características rectangulares tais como a matrícula, o vidro, o pára-choques, etc. É possível verificar então que a detecção dos cantos utilizando os resultados da detecção de contornos com os gradientes é mais eficaz que utilizando o algoritmo de Sobel.

Na figura 4.38 é representado um processo similar, usando também o resultado da extracção dos contornos com o algoritmo de Sobel e o cálculo dos gradientes, mas desta vez ao contrário do processo anterior irá ser feita uma união dos resultados verticais e horizontais para tentar encontrar a existência de formas rectangulares.



Figura 4.38: Extracção de formas rectangulares utilizando a união dos contornos verticais e horizontais, calculados com o algoritmo de Sobel (imagem da esquerda) e cálculo dos gradientes (imagem da direita)

O processo representado na figura 4.38 demonstra bons resultados, extraindo as características mais rectangulares dos veículos. A primeira imagem da figura 4.38 continua a conter algum ruído, resultante dos limites laterais das faixas. Para detecção destas características é possível a utilização de modelos pré-definidos da geometria de um veículo e através de técnicas de template-matching (ver secção 3.5.2) para tentar encontrar possíveis localizações de veículos que correspondam tanto quanto possível ao modelo.

Para conseguir atingir algum sucesso com esta técnica seria necessário algum processamento prévio para eliminar algum ruído, bem como o conhecimento da transformação aplicada pelo efeito perspectiva, por forma a conseguir validar alguns rectângulos detectados pelo seu tamanho de acordo com o sítio onde aparecem na imagem. Como é óbvio um objecto que aparece próximo do topo da imagem tem que ter uma largura e altura máxima muito inferior a um objecto que apareça próximo do fundo da imagem. É possível também eliminar alguns resultados errados utilizando a relação esperada entre a largura e a altura de um veículo, uma vez que normalmente esta relação converge para um determinado valor, excepção feita aos veículos anormalmente altos como será o caso de alguns veículos pesados.

Proposta de solução

Tal como referido no início desta secção, a generalidade dos sistemas referidos na literatura que utilizam as características comuns nos veículos para fazer a sua detecção, seguem uma metodologia de geração e verificação de hipóteses (ver figura 4.27). A solução aqui proposta para efectuar a detecção de veículos funciona também com este tipo de metodologia, dividindo-se assim em duas tarefas principais: geração de hipóteses da localização de possíveis veículos e a sua posterior verificação.

Para execução destas duas tarefas irão ser utilizadas as características que se concluíram ser mais importantes, ou que permitiram obter uma extracção com melhores resultados, tendo em conta o que foi estudado até agora nesta secção. A análise feita até então permite concluir que os contornos e simetrias verticais e horizontais, juntamente com a sombra, são as características que mais facilmente podem contribuir para uma eficiente detecção de veículos.

Para a tarefa de geração de hipóteses é necessário um método que consiga gerar um grupo de hipóteses em que a localização de todos os veículos reais estejam presentes nesse grupo, embora o processo resulte também na detecção de várias hipóteses erradas. Neste caso o processo proposto para gerar as várias hipóteses utiliza a sombra e os contornos horizontais da imagem, uma vez que são estes elementos que melhor evidenciam a presença de veículos.

Verificou-se experimentalmente que a intersecção entre os resultados da extracção das características anteriores permitem por si só gerar resultados bastante promissores para a detecção de veículos. Várias técnicas que utilizam metodologias semelhantes a esta são descritas em [67, 68, 62, 69].

As imagens na figura 4.39 representam os dois ambientes onde irão ser aplicados os procedimentos de detecção.

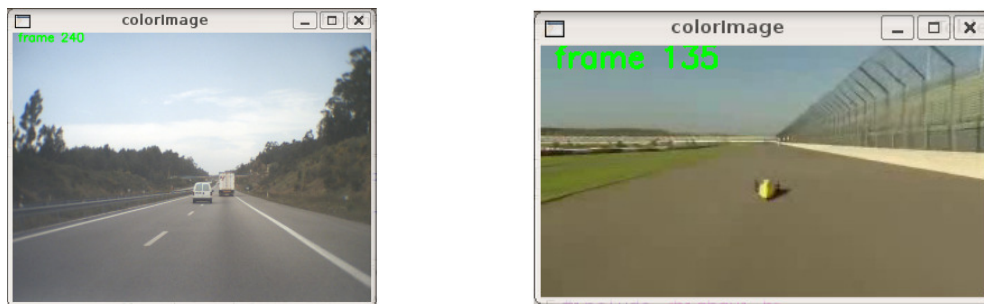


Figura 4.39: Imagens sobre as quais irá ser aplicado o processo de detecção de veículos.

O passo inicial deste processo é fazer uma extracção das sombras e dos contornos horizontais. O processo de extracção de sombras, já referido anteriormente nesta secção, apresentava um problema que era a detecção de bastante ruído principalmente nas zonas superiores da imagem, correspondente ao ambiente circundante. Para minimizar esta situação, este processo foi ligeiramente alterado. Uma vez que a detecção de sombra é feita analisando a imagem de baixo para cima, apenas a primeira ocorrência de sombra é considerada, ignorando assim as restantes sombras que poderiam ser detectadas em partes mais superiores da imagem, que não são interessantes para o processo. Desta forma é possível fazer a detecção apenas das sombras que são detectadas na parte mais inferior da imagem, o que inclui as sombras dos veículos, ignorando assim os restantes resultados que não são interessantes para o processo.

Os contornos horizontais são extraídos com o algoritmo de Sobel, uma vez que este foi o método que se mostrou mais eficiente. De seguida é aplicado a ambos os resultados da extracção de sombra e contornos horizontais um processo de dilatação para realçar as características encontradas. O resultado desta operação é mostrado na figura 4.40.

A intersecção das imagens resultantes é então usada para o passo inicial da geração de hipóteses da localização dos veículos. Como é possível verificar nas imagens do lado direito das figura 4.40, são detectados os contornos inferiores dos veículos, embora estejam também presentes características que correspondem aos limites laterais das faixas.

De seguida é utilizado um processo para detecção de segmentos horizontais presentes na imagem, com o objectivo de localizar as características mais relevantes na imagem, que podem representar os

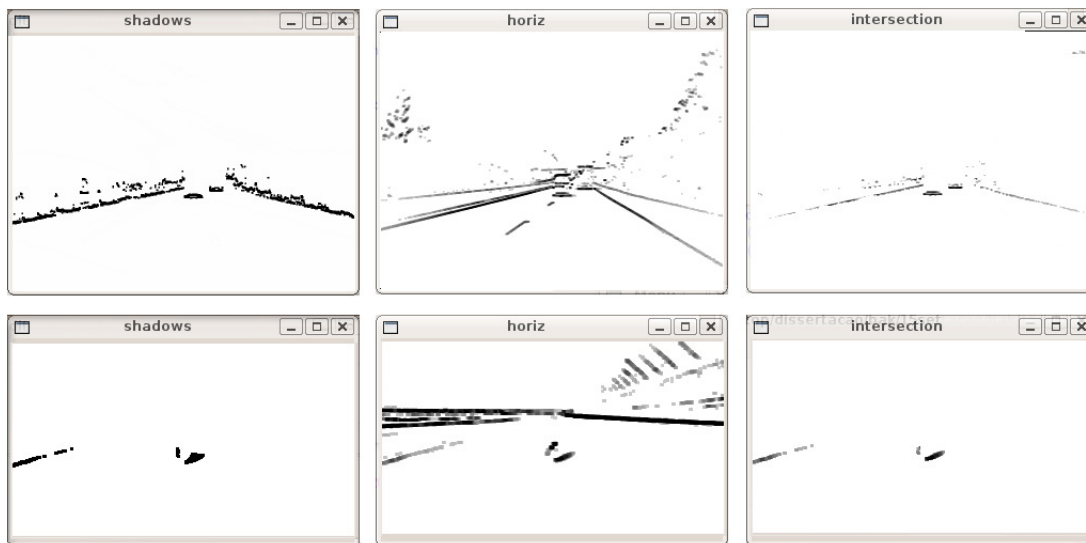


Figura 4.40: Da esquerda para a direita: extracção de sombras, extracção de contornos horizontais e o resultado da intersecção entre as duas imagens anteriores.

veículos. O resultado deste processo é apresentado nas imagens da esquerda da figura 4.41.

Tal como é possível verificar nestas imagens, existe um elevado número de segmentos detectados, existindo situações em que vários segmentos pertencem ao mesmo objecto. Assim, por forma a simplificar a verificação destas hipóteses é efectuado um processo de agrupamento de segmentos horizontais que estão localizadas em sítios muito próximos, e que provavelmente pertencem a um só objecto (ver imagens intermédias da figura 4.41).

De seguida são geradas finalmente as hipóteses da localização dos veículos. Os segmentos detectados são utilizados para construir uma região de análise imediatamente em cima destes, de forma quadrada, uma vez que esta forma consegue representar de forma eficiente a generalidade dos veículos. O resultado final do processo de geração de hipóteses é representado pelas imagens do lado direito da figura 4.41.

Está assim completo o processo de geração de hipóteses, sendo agora necessário de alguma forma fazer uma validação destas mesmas hipóteses para verificar quais é que mais provavelmente representam veículos. A forma mais natural para o fazer é a utilização das características anteriormente mencionadas, simetria e contornos, que se encontram presentes nos veículos como em nenhum outro lado da imagem.

A solução encontrada foi juntar os resultados da extracção vertical e horizontal de contornos e simetria, numa única imagem, por forma a obter uma peça única de análise que contenha informação sobre todas estas características. O resultado deste processo é demonstrado pelas imagens do lado esquerdo da figura 4.42. A estas imagens iremos chamar imagens de energia, de uma determinada cena, para facilitar a referência ao resultado deste procedimento. É possível verificar que nestas imagens as localizações dos veículos encontram-se particularmente bem representadas, tal como seria de esperar.

O processo de validação das hipóteses deverá ser dividido em duas tarefas principais:

- Verificação em cada frame da probabilidade de cada hipótese representar um veículo real.
- Verificação da ocorrência de cada hipótese em frames consecutivas, ao longo da sequência de frames.

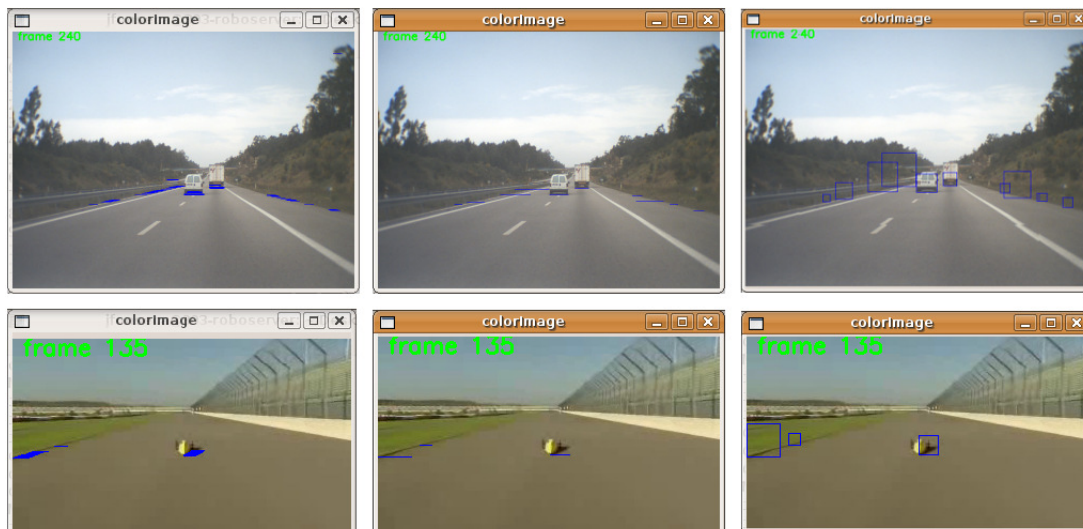


Figura 4.41: Da esquerda para a direita: detecção de segmentos horizontais, agrupamento de segmentos horizontais semelhantes e a geração final de hipóteses.

A primeira tarefa utiliza a imagem energia (ver imagens do lado esquerdo da figura 4.42), calculando a média aritmética dessa imagem em cada zona correspondente a cada hipótese, atribuindo assim uma determinada energia a cada hipótese. Um valor elevado de energia significa que essa hipótese é constituída por um número elevado de características, contornos e simetria, ou seja, quanto maior energia tiver uma hipótese maior é a probabilidade de esta representar um veículo. Utilizando as energias de cada hipótese, estas são classificadas como válidas ou não válidas (ver imagens do meio na figura 4.42). Este processo resulta na validação de várias hipóteses, normalmente incluindo algumas falsas detecções, o que confirma a incapacidade de um método de detecção utilizar apenas uma frame.

A segunda tarefa do processo de validação é verificar se as hipóteses consideradas válidas pelo processo anterior, se encontram na frame imediatamente a seguir. Caso uma hipótese considerada válida se encontre presente num razoável número de frames consecutivas é classificada como sendo um possível veículo (ver imagens do lado direito da figura 4.42).

Executando este processo ao longo de uma sequência de imagens, é possível apresentar com bom nível de confiança hipóteses válidas da localização de veículos reais.

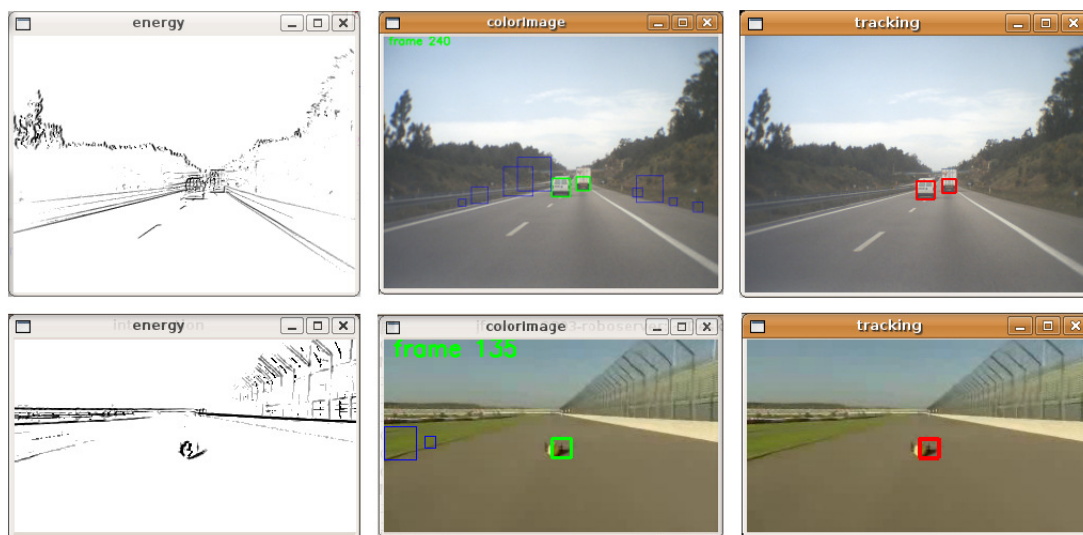


Figura 4.42: Da esquerda para a direita: cálculo da imagem energia, verificação das hipóteses numa frame, verificação das hipóteses na sequência de frames.

Capítulo 5

Avaliação crítica de resultados

Neste capítulo irá ser analisada de forma geral a eficiência dos processos de detecção da faixa de rodagem e de detecção dos seus limites laterais, descritos nas secções 4.4 e 4.5, focando as situações críticas no funcionamento destes. Por situações críticas entendem-se situações que alteram o normal funcionamento dos mesmos, como por exemplo mudanças bruscas no padrão da estrada, presença de grandes zonas afectadas por sombra ou a presença de outros veículos.

A secção relativa à detecção de obstáculos irá igualmente ser alvo de discussão, avaliando o resultado do estudo efectuado com o objectivo de perceber quais são as características comuns nos veículos que podem ser mais proveitosas para fazer a sua detecção. De seguida irá ser também comentado o mecanismo de detecção de veículos elaborado com base neste mesmo estudo, tendo em conta os seus pontos críticos.

5.1 Detecção da faixa de rodagem

O processo de detecção da faixa de rodagem tem como ponto de partida uma imagem sem qualquer tipo de processamento, tendo como objectivo classificar as várias zonas desta mesma imagem como estrada (zonas a branco) ou não estrada (zonas a preto). As várias etapas deste processo podem ser resumidas pelo diagrama representado na figura 5.1.

O sistema revela boa capacidade de adaptação em situações de mudanças bruscas no padrão da estrada, uma vez que o padrão válido da estrada é calculado frame a frame, criando uma independência do resultado deste processo nas frames anteriores. No entanto, no momento exacto de transição no padrão da estrada o processo de segmentação é pouco eficiente uma vez que os sensores padrão podem estar situados em dois padrões diferentes da estrada simultaneamente (ver figura 5.2).

Uma situação semelhante a esta mudança brusca de padrão da estrada é a passagem em zonas em que a estrada é fortemente afectada por sombra. No entanto a passagem em zonas com sombra é um acontecimento que pode ser detectado uma vez que a sombra pode ser identificada através de características próprias (ver secção 4.3).

As zonas fortemente afectadas por sombra podem ser detectadas, minimizando o seu impacto nas tarefas seguintes, tal como na detecção dos limites laterais das faixas de rodagem. No entanto, tal como tinha sido referido anteriormente, quando os sensores se situam simultaneamente em regiões com diferentes padrões, como é o caso de zonas com e sem sombra, o processo de detecção da faixa de rodagem apresenta maus resultados (ver figura 5.3).

Nas situações em que outros veículos circulam também na estrada, a zona onde estes veículos se situam irá ser classificada como não estrada, uma vez que os veículos normalmente apresentam

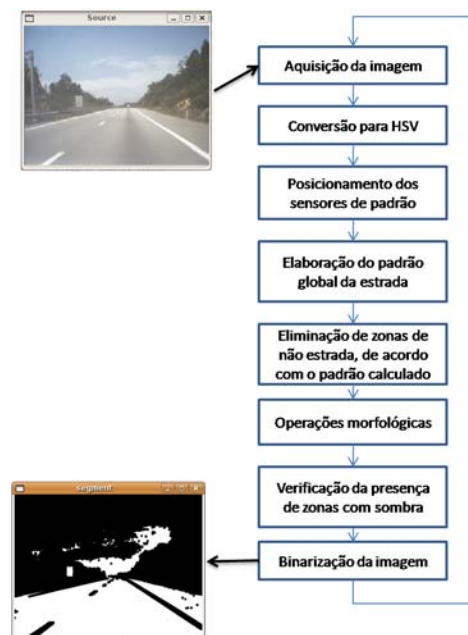


Figura 5.1: Processo de detecção dos limites laterais da faixa de rodagem.

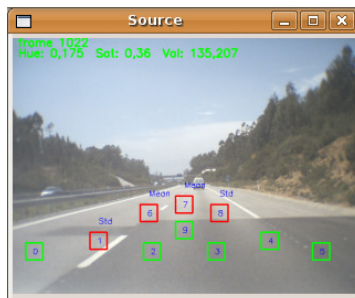


Figura 5.2: Impacto da mudança brusca do padrão da estrada no processo de detecção da faixa de rodagem.

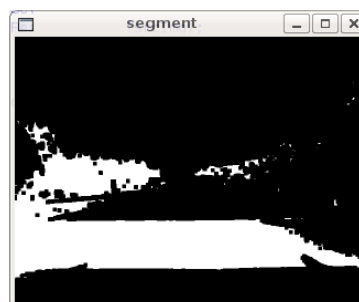
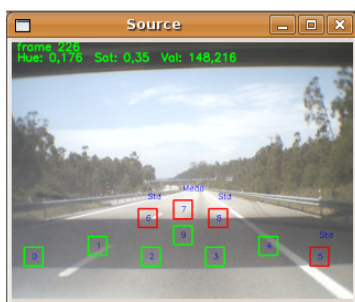


Figura 5.3: Impacto das sombras no processo de detecção da faixa de rodagem.

características bem distintas da faixa de rodagem (ver figura 5.4).

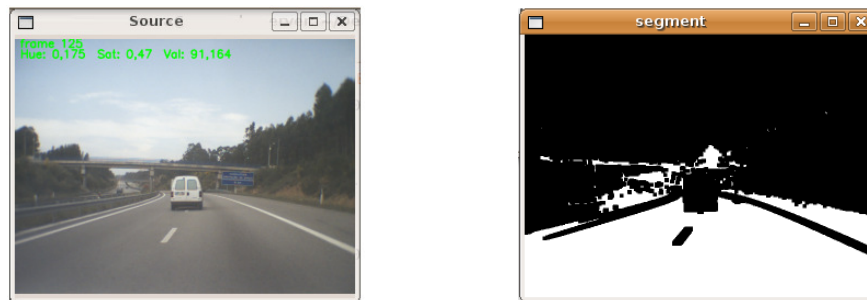


Figura 5.4: Impacto do aparecimento de veículos no processo de segmentação.

As linhas brancas, caso existam, serão também classificadas como não estrada, uma vez que são constituídas por características muito diferentes da região da estrada. No entanto era este o comportamento desejável, uma vez que no caso de o ambiente ser estruturado são estas faixas brancas que representam os limites laterais das faixas de rodagem.

Um dos requisitos principais deste processo é a capacidade de funcionar em ambientes estruturados ou não estruturados da mesma forma, sendo tão adaptável e eficiente quando possível. Em ambientes não estruturados o processo de segmentação apresenta um bom comportamento, caso exista um padrão dominante que defina a zona de estrada. Em ambientes estruturados, uma vez que as faixas brancas são também segmentadas, o processo consegue de forma igualmente eficaz efectuar a classificação das zonas da estrada.

Pode então ser concluído que os resultados obtidos em situações normais através deste processo são satisfatórios, conseguindo atingir o objectivo a que se propões na generalidade das situações. No entanto, em situações de mudança brusca no padrão da estrada, o processo apresenta algumas limitações.

5.2 Detecção dos limites da faixa de rodagem

As etapas do processo de detecção dos limites da faixa de rodagem pode ser descrito de forma resumida pelo diagrama representado na figura 5.5. Este processo tem como ponto de partida o resultado do processo de detecção da faixa de rodagem, uma imagem binária, representando a branco as zonas de estrada e a preto as zonas de não estrada.

Em ambientes estruturados, onde existe a presença de faixas brancas delimitadoras, considera-se que os limites laterais das faixas de rodagem são representados pelas próprias faixas brancas. Em ambientes não estruturados os limites laterais são representados pela transição entre a zona de estrada e as zonas laterais que a delimitam, qualquer que seja a constituição destas mesmas zonas.

Uma situação que prejudica o normal funcionamento deste processo é a presença de outros veículos na estrada, que provocam a oclusão total ou parcial dos limites laterais da estrada, sendo nesta situação impossível detectar com eficiência estes mesmos limites (ver figura 5.6).

Em zonas afectadas por sombra o processo de detecção da faixa de rodagem apresenta-se pouco eficiente. Consegue-se, no entanto, fazer a detecção deste tipo de zonas por forma a minimizar o seu impacto nesta tarefa de detecção dos limites laterais. Quando é detectada uma zona fortemente afectada por sombra o processo de detecção de limites laterais interrompe o seu normal funcionamento e estima a sua posição com base na última posição detectada (ver figura 5.7).

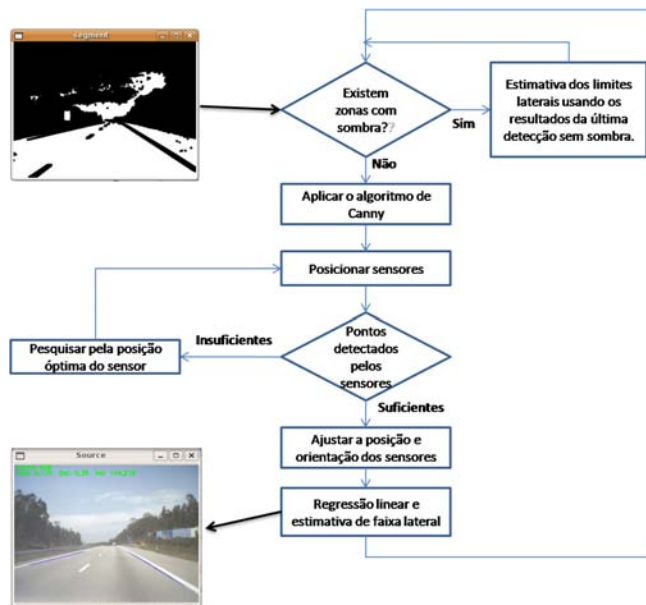


Figura 5.5: Processo de detecção dos limites laterais da faixa de rodagem.

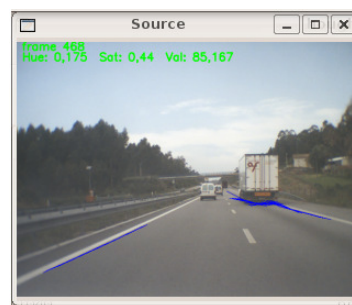
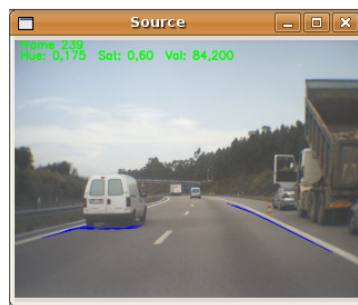


Figura 5.6: Impacto do aparecimento de veículos no processo de detecção dos limites laterais da faixa.



Figura 5.7: Impacto da sombra no processo de detecção dos limites laterais da faixa.

Esta estimativa da localização dos limites das faixas normalmente não se afasta muito da realidade a menos que a zona de sombra se estenda ou que a geometria da estrada sofra uma brusca mudança. Após a passagem deste tipo de zonas, o processo retoma o seu normal funcionamento.

Um outro factor que pode causar problemas no processo de detecção dos limites laterais é a possibilidade de mudança brusca na localização desses limites ou de um mau resultado do processo de detecção da faixa de rodagem. Ambas situações fazem com que os sensores utilizados para efectuar a detecção não consigam obter nenhum resultado, perdendo a localização dos limites laterais.

Para ultrapassar este problema foi desenvolvido um mecanismo que permite aos sensores procurarem automaticamente a sua melhor posição, sempre que estes deixarem de detectar qualquer contorno, ou seja, quando perdem o limite lateral da estrada. Este mecanismo consiste numa pesquisa horizontal e vertical, nas zonas onde provavelmente se situam os limites laterais, de forma a encontrar uma posição óptima que se caracteriza neste caso pelo máximo de pontos encontrados.

Além deste mecanismo de procura da melhor posição, o sensor adapta-se também automaticamente à medida que os limites laterais se deslocam, devido à movimentação do veículo. Esta adaptação é feita nos eixos X e Y bem como na inclinação do sensor, adaptando-se da melhor forma à geometria dos limites laterais das faixas. Desta forma é possível ultrapassar as variações das posições dos limites laterais, uma vez que os sensores se adaptam dinamicamente à melhor posição em cada momento.

De forma geral este processo apresenta-se com bom comportamento em situações normais, conseguindo detectar com sucesso os limites laterais da faixa de rodagem.

5.3 Detecção de obstáculos

Uma vez que o processo de detecção de veículos proposto tem como base as características comuns dos veículos, foi inicialmente efectuado um estudo analisando as várias características por forma a verificar quais seriam as mais vantajosas. Tal como já foi anteriormente referido, os contornos verticais/horizontais, a simetria, a cor e textura, a sombra, as luzes dos veículos e a sua forma rectangular são as características principais que normalmente são utilizadas para tarefas de detecção.

Da análise feita na secção 4.6 podemos concluir que as características mais interessantes para efectuar a detecção são os contornos horizontais e verticais e a sombra. Estas características revelam os melhores resultados na tarefa de identificação da posição dos veículos, sendo por isso características de eleição para elaborar o processo de geração de hipóteses das localizações dos veículos.

A extracção dos contornos horizontais e verticais resulta normalmente numa boa indicação da posição do(s) veículo(s), uma vez que estes são objectos que revelam tantas características verticais e horizontais como nenhum outro num ambiente de estrada. A sombra pelas suas características muito próprias, torna possível a sua extracção da imagem em situações de normal iluminação, revelando-se também uma característica bastante importante para identificar a posição do veículo. Das restantes características, a simetria é a que revela ainda resultados satisfatórios na identificação de um veículo. No entanto, o cálculo da simetria pode produzir por vezes falsas detecções uma vez que o ambiente circundante cria por vezes razões de simetria, devendo esta característica ser utilizada como forma complementar e não como primeira forma de detecção. Assim a utilização da simetria é mais indicada no processo validação das hipóteses geradas, para classificar estas hipóteses como válidas ou não válidas.

As restantes características referidas não produzem resultados de relevância, tal como já foi comentado na secção 4.6, motivo pelo qual se conclui que a sua utilização como método de detecção não será muito proveitosa.

Tendo em conta o resultado do estudo efectuado sobre as várias características, foi desenvolvido um mecanismo de detecção de veículos. O processo de detecção de veículos proposto tenta tanto quanto possível respeitar a metodologia de geração e verificação de hipóteses, usando para isso as características que se revelam mais comuns na generalidade dos veículos. Para que uma hipótese de um veículo seja considerada válida é necessário que esta apresente características e comportamentos próprios de um veículo. O diagrama apresentado na figura 5.8 representa de forma resumida todo o processo da solução proposta para detecção de veículos.



Figura 5.8: Diagrama exemplificativo do processo de detecção de veículos.

Este mecanismo de detecção de veículos, uma vez que não se baseia num setup específico de aquisição de imagem, foi desenvolvido por forma a ser tão genérico quanto possível. Desta forma é possível alterar no mecanismo vários parâmetros que alterem o seu método de funcionamento por forma a ser possível a sua adaptação a uma ambiente mais específico.

A figura 5.9 mostra duas frames da mesma sequência de vídeo, e como é possível verificar existe um intervalo de tempo desde o momento em que o veículo da direita fica bem visível (na imagem do lado esquerdo) até que seja considerado uma hipótese válida da localização de um veículo (imagem do lado direito). Este intervalo de tempo é devido ao processo de validação das hipóteses encontradas. Na imagem do lado esquerdo é demonstrado aproximadamente o momento em que é criada uma hipótese de detecção referente ao veículo da direita, na imagem da direita é o momento em que esta mesma hipótese é validada. Este intervalo de tempo é um parâmetro do sistema, podendo ser alterado por forma a melhor cumprir os objectivos, dependendo da situação a que se destina. Por exemplo, se se trata de um ambiente em que os veículos circulam a uma velocidade moderada é possível aumentar o intervalo de tempo de validação por forma a validar as hipóteses com mais certeza. No entanto se os veículos circularem com maiores velocidades este intervalo de tempo pode tornar-se crítico por não ser possível detectar os veículos em tempo útil.



Figura 5.9: Intervalo de tempo necessário para a validação de uma hipótese de localização de um veículo.

Um outro parâmetro do sistema é o tamanho mínimo a partir do qual é possível gerar uma hipótese da localização de um veículo. Tal como no caso anterior, este parâmetro depende do ambiente para o qual o sistema se destina a funcionar e do tamanho dos veículos que se espera encontrar. Na imagem do lado direito da figura 5.10 é possível verificar que é detectado um veículo a uma distância considerável, uma vez que a distância mínima para geração de hipóteses é pequena. No entanto na imagem do lado esquerdo da mesma figura é possível mostrar que várias outras hipóteses de tamanho reduzido são também inicialmente detectadas, devido a este parâmetro ter um valor tão reduzido.



Figura 5.10: Processo de detecção de veículos.

Este mecanismo de detecção de veículos revela-se eficiente em situações normais, existindo no entanto várias situações em que os resultados apresentados não são os desejáveis. Na figura 5.11 são mostradas situações de mau funcionamento deste mecanismo. Na primeira situação, imagem da esquerda da figura 5.11, 2 veículos são considerados apenas como um só, devido à sua proximidade. Na segunda situação, imagem do meio da figura 5.11, é mostrado um exemplo de falsa detecção. O sinal de trânsito é considerado um veículo, uma vez que revela fortes características de simetria e contornos, tal como é esperado de um veículo. Na terceira situação, imagem da direita da figura 5.11, é detectado um veículo mas a sua proximidade ao limite lateral da estrada faz com que a detecção seja feita de forma errada, detectando um tamanho superior ao real.

O problema da falsa detecção do sinal de trânsito pode ser resolvido caso exista um mecanismo complementar de detecção da zona de estrada, eliminando assim todas as hipóteses localizadas fora da mesma. Os restantes dois problemas são possíveis de ultrapassar se a transformação na imagem devido ao efeito perspectiva for conhecida. Desta forma é possível invalidar falsas detecções, comparando o tamanho das hipóteses geradas com o tamanho expectável de um veículo em cada zona da imagem. A implementação de um mecanismo que elimine a zona superior à linha do horizonte na imagem, eliminando assim possibilidade de geração incorrecta de hipóteses a cima deste nível, pode também aumentar a eficiência deste processo.



Figura 5.11: Detecções erradas de veículos.

Tal como já foi referido, este trabalho não tem um setup físico de aquisição de imagem definido não sendo assim possível construir um modelo que permita conhecer a transformação na imagem devido ao efeito perspectiva. A construção deste modelo revela-se muito importante na tarefa de detecção de veículos uma vez que permite ter a noção da relação entre cada porção da imagem e o que esta representa no ambiente real. Desta forma o processo proposto não inclui nenhuma validação feita através deste conhecimento.

De forma geral, mesmo tendo em conta as situações atrás referidas, é possível dizer que o processo apresenta uma eficiência relativamente satisfatória. Uma vez conhecido o setup final de aquisição, e o ambiente para o qual se destina actuar, é possível introduzir no processo várias condicionantes que irão certamente melhorar a sua eficiência.

Capítulo 6

Conclusões e trabalho futuro

Nesta secção irão ser feitas algumas conclusões gerais sobre o trabalho efectuado, incidindo sobre as suas três principais categorias: detecção da faixa de rodagem, detecção dos limites laterais e detecção de obstáculos. Serão também comentados vários aspectos que poderiam ser melhorados, de forma a evoluir para um sistema mais eficiente.

O objectivo inicial de desenvolvimento de um sistema de detecção da faixa de rodagem, que funcionasse em diferentes tipos de ambiente, estruturados e não estruturados, foi alcançado. O sistema é eficiente, conseguindo, na generalidade das situações, uma correcta classificação de regiões de imagem em estrada e não estrada, facilitando o trabalho para o processo que se segue, a detecção dos limites laterais da faixa de rodagem.

Identificam-se, no entanto, situações em que é possível melhorar. A adaptação automática das regiões de análise, de acordo com a geometria da estrada em cada momento ou com o posicionamento do veículo na estrada, é uma delas. Numa situação em que o veículo se desloca para a esquerda ou para a direita a localização das regiões de análise devia igualmente acompanhar este movimento, por forma a garantir que tantas regiões quanto possível se mantivessem localizadas na faixa de rodagem e não fora dela. Na figura 6.1 ilustra-se uma situação onde, em consequência do movimento do veículo para a esquerda, as regiões de análise estão colocadas fora da faixa de rodagem.

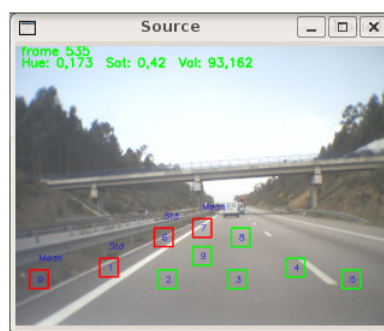


Figura 6.1: Exemplo de mau posicionamento das regiões de análise.

O conhecimento prévio da localização do sistema de aquisição de imagem também permite melhorar o sistema de detecção. Esse conhecimento permite perceber melhor o que cada porção de imagem representa no ambiente real, sendo assim possível desenvolver um mecanismo de auto-localização das regiões de análise.

Como conclusão é possível dizer que a solução encontrada, que consiste em elaborar um perfil

cromático da estrada usando sensores padrão, demonstrou-se ser uma boa aposta com bons resultados obtidos. No entanto existem problemas que é necessário resolver, tal como é o caso das situações com bruscas transições de padrão da estrada.

O processo de detecção dos limites laterais das faixas depende directamente do processo anterior, uma vez que a imagem resultante do processo de detecção da faixa de rodagem é o ponto de partida da tarefa de detecção dos limites laterais. A solução encontrada para fazer a detecção dos limites laterais com custos computacionais aceitáveis foi a inclusão de sensores de linha.

Estes sensores revelam-se bastante eficazes no seu objectivo, conseguindo detectar os limites laterais quando se encontram bem posicionados. No entanto, existe a necessidade de posicionar correctamente os sensores por forma a que estes se encontrem numa posição próxima dos limites, bem como alterar esta posição e a sua orientação de acordo com as mudanças geométricas da estrada em cada momento. Para estas necessidades os mecanismos adequados foram criados, conseguindo assim lidar da melhor forma com o problema.

Existem também outros problemas tais como a oclusão dos limites laterais por parte de outros veículos que torna impossível a sua correcta detecção nestas situações específicas.

No futuro, a criação de uma estrutura que representasse o estado do sistema, podendo de alguma forma salvaguardar o resultado da detecção de uma forma progressiva, seria provavelmente muito proveitoso para eliminação de detecções erradas. Conjuntamente com a utilização de filtros de Kalman [70, 71], a história de detecção dos limites laterais nas frames anteriores poderia ajudar na previsão da localização dos limites laterais nas frames seguintes. Desta forma, quando não fosse possível por algum motivo detectar os limites laterais, como no caso da sua oclusão por parte de veículos, uma estimativa da sua localização poderia ser feita recorrendo a um mecanismo semelhante ao mencionado.

Quanto à detecção de obstáculos, o trabalho efectuado teve como objectivo inicial o estudo de técnicas de extracção das características comuns na generalidade dos veículos, por forma a verificar quais seriam as mais proveitosas para efectuar a sua detecção. Foi então possível concluir que as características dos veículos que são mais relevantes são os seus contornos e a própria sombra. Estas características estão sempre presentes nos veículos e revelam-se relativamente fáceis de extrair, excepto em situações de mudanças bruscas de iluminação que podem deteriorar o resultado da extracção destas características.

Após o estudo anterior um sistema de detecção foi proposto. Devido ao desconhecimento do sistema específico de aquisição e do ambiente em que destina a actuar, é difícil construir um sistema que possa ser tão geral e adaptável por forma a funcionar com qualquer sistema de aquisição e que detecte quaisquer tipo de veículos. No entanto o sistema apresenta já um nível de eficiência razoável, sendo certamente possível melhorar se forem conhecidos os aspectos anteriormente referidos.

Outro factor que aumenta a falibilidade do processo é que ainda não foi feita a integração do mecanismo de segmentação e detecção dos limites laterais com o sistema de detecção de obstáculos. Esta integração será também certamente proveitosa uma vez que será possível fazer a análise apenas em zona de estrada, eliminando características irrelevantes da imagem como é o caso do ambiente circundante à faixa de rodagem. Para fazer esta integração é no entanto necessário aumentar a robustez dos mecanismos referidos, detecção da faixa de rodagem e detecção dos limites laterais, por forma a que estes se tornem ainda mais fiáveis e eficientes.

Bibliografia

- [1] Icaro website . <http://icaro.ua.pt/>, consultado em Julho de 2009.
- [2] Shell Eco-marathon . www.shell.com/ecomarathon/, consultado em Julho de 2009.
- [3] Ricardo Ângelo dos Santos Filipe. Teleoperação de um veículo da eco-maratona. Master's thesis, Universidade de Aveiro, 2008.
- [4] L. Figueiredo, I. Jesus, J.A.T. Machado, J.R. Ferreira, and J.L. Martins de Carvalho. Towards the development of intelligent transportation systems. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pages 1206–1211, 2001.
- [5] B. Ulmer. Vita-an autonomous road vehicle (arv) for collision avoidance in traffic. In *Intelligent Vehicles '92 Symposium., Proceedings of the*, pages 36–41, Jun-1 Jul 1992.
- [6] R. Bishop. A survey of intelligent vehicle applications worldwide. In *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE*, pages 25–30, 2000.
- [7] Massimo Bertozzi, Alberto Broggi, Alessandra Fascioli, and Ra Fascioli A. Vision-based intelligent vehicles: State of the art and perspectives, 2000.
- [8] C. H. Chen, L. F. Pau, and P. S. P. Wang, editors. *Handbook of Pattern Recognition and Computer Vision*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2000.
- [9] E.D. Dickmanns, R. Behringer, C. Brudigam, D. Dickmanns, F. Thomanek, and V. van Holt. An all-transputer visual autobahn-autopilot/copilot. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*, pages 608–615, May 1993.
- [10] R. Behringer, V. V. Holt, and D. Dickmanns. Road and relative ego-state recognition. In *Intelligent Vehicles '92 Symposium., Proceedings of the*, pages 385–390, Jun-1 Jul 1992.
- [11] R. Behringer and R.B.M. Maurer. Results on visual road recognition for road vehicle guidance. In *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*, pages 415–420, Sep 1996.
- [12] E.D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek, and J. Schiehlen. The seeing passenger car 'vamors-p'. pages 68–73, Oct. 1994.
- [13] Ernst Dieter Dickmanns and Volker Graefe. Dynamic monocular machine vision. *Mach. Vision Appl.*, 1(4):223–240, 1988.
- [14] Ernst Dieter Dickmanns. Machine Perception Exploiting High-Level Spatio-Temporal Models. *AGARD LS 185, Machine Perception, Hampton, VA, USA; Munich; Madrid*, 1992.

- [15] Carnegie Mellon University website . <http://www.ri.cmu.edu/>, consultado em Julho de 2009.
- [16] Dirk Langer Julio Rosenblatt Dean Pomerleau, Chuck Thorpe and Rahul Sukthankar. Avcs research at carnegie mellon university. In *Proceedings of Intelligent Vehicle Highway Systems*, 1994.
- [17] C. Thorpe, M.H. Hebert, T. Kanade, and S.A. Shafer. Vision and navigation for the carnegie-mellon navlab. 10(3):362–373, May 1988.
- [18] H. T. Kung. Warp experience: we can map computations onto a parallel computer efficiently. In *ICS '88: Proceedings of the 2nd international conference on Supercomputing*, pages 668–675, New York, NY, USA, 1988. ACM.
- [19] J.D. Crisman and J.A. Webb. The warp machine on navlab. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(5):451–465, May 1991.
- [20] Sanjiv Singh D. Feng and Bruce Krogh. Implementation of dynamic obstacle avoidance on the cmu navlab. In *Proc. IEEE International Conference on Systems Engineering*, August 1990.
- [21] M. Bertozzi, A. Broggi, M. Cellario, A. Fascioli, P. Lombardi, and M. Porta. Artificial vision in road vehicles. *Proceedings of the IEEE*, 90(7):1258–1271, Jul 2002.
- [22] D. Bishop. Vehicle-highway automation activities in the united states. In *Proceedings of the International AHS Workshop, US Department of Transportation*, 1997.
- [23] S. Tsugawa. Vision-based vehicles in japan: machine vision systems and driving control systems. *Industrial Electronics, IEEE Transactions on*, 41(4):398–405, Aug 1994.
- [24] H. Tokuyama. Asia-pacific projects status and plans. In *Proceedings of the International AHS Workshop, US Department of Transportation*, 1997.
- [25] Advanced Cruise-Assist Highway System Research Associationr . <http://www.ri.cmu.edu/>, consultado em Julho de 2009.
- [26] Defense Advanced Research Projects Agency website. <http://www.darpa.mil>, consultado em Junho de 2009.
- [27] E. Jones, B. Fulkerson, E. Frazzoli, D. Kumar, R. Walters, J. Radford, and R. Mason. Autonomous off-road driving in the darpa grand challenge. In *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, pages 366–371, 25-27, 2006.
- [28] Carnegie Mellon Tartan Racing website . <http://www.tartanracing.org/>, consultado em Julho de 2009.
- [29] Chris Urmson, Joshua Anhalt, Drew Bagnell, Christopher Baker, Robert Bittner, M. N. Clark, John Dolan, Dave Duggins, Tugrul Galatali, Chris Geyer, Michele Gittleman, Sam Harbaugh, Martial Hebert, Thomas M. Howard, Sascha Kolski, Alonzo Kelly, Maxim Likhachev, Matt McNaughton, Nick Miller, Kevin Peterson, Brian Pilnick, Raj Rajkumar, Paul Rybski, Bryan Salesky, Young-Woo Seo, Sanjiv Singh, Jarrod Snider, Anthony Stentz, William “Red” Whitaker, Ziv Wolkowicki, Jason Ziglar, Hong Bae, Thomas Brown, Daniel Demitrish, Bakhtiar Litkouhi, Jim Nickolaou, Varsha Sadekar, Wende Zhang, Joshua Struble, Michael Taylor,

- Michael Darms, and Dave Ferguson. Autonomous driving in urban environments: Boss and the urban challenge. *J. Field Robot.*, 25(8):425–466, 2008.
- [30] Sociedade Portuguesa de Robotica - SPR. <http://www.spr.ua.pt/>, consultado em Setembro de 2009.
- [31] José Nuno da Silva Carvalho. Rota 2008: um robô para condução autónoma. Master’s thesis, Universidade de Aveiro, 2008.
- [32] Critical Move web page . <http://www.critical-move.com/>, consultado em Setembro de 2009.
- [33] Critical Challenge web page . <http://www.criticalchallenge.com/>, consultado em Setembro de 2009.
- [34] CyberMove web page . <http://www.cybermove.org/>, consultado em Setembro de 2009.
- [35] Laboratório de Robótica Móvel web page. <http://paloma.isr.uc.pt/mrl/index.php>, consultado em Setembro de 2009.
- [36] Autonomous Systems Laboratory (LSA) home page . <http://www.lsa.isep.ipp.pt/>, consultado em Agosto de 2009.
- [37] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection using optical sensors: a review. In *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*, pages 585–590, Oct. 2004.
- [38] Michael Shneier Tommy, Tommy Chang, Tsai Hong, Gerry Cheok, Harry Scott, Steve Legowik, and Alan Lytle. A repository of sensor data for autonomous driving research. In *in Proceedings of the SPIE Aerosense Conference*, 2003.
- [39] Mario Bellino Sascha Kolski, David Ferguson and Roland Siegwart. Autonomous driving in structured and unstructured environments. In *IEEE Intelligent Vehicles Symposium*, 2006.
- [40] M. Bertozzi and A. Broggi. Gold: a parallel real-time stereo vision system for generic obstacle and lane detection. *Image Processing, IEEE Transactions on*, 7(1):62–81, Jan 1998.
- [41] Davide Scaramuzza and Roland Siegwart. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *IEEE Transactions on Robotics*, 24(5):1015–1026, 2008.
- [42] M. Tkalcic and J.F. Tasic. Colour spaces: perceptual, historical and applicational background. In *EUROCON 2003. Computer as a Tool. The IEEE Region 8*, volume 1, pages 304–308 vol.1, Sept. 2003.
- [43] Luís Miguel de Oliveira Matos. Estudo e aplicações da norma de codificação de video h.264/avc. Master’s thesis, Universidade de Aveiro, 2009.
- [44] The home page of Armando J. Pinho. <http://www.ieeta.pt/~ap/>, consultado em Junho de 2009.
- [45] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O’Reilly Media, Inc., 1st edition, October 2008.

- [46] Li Zhang, Wen-Hui Zhou, and Ji-Lin Liu. A robust road segmentation method. In *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, volume 2, pages 912–917, Nov. 2007.
- [47] V. Graefe. Dynamic vision systems for autonomous mobile robots. In *Intelligent Robots and Systems '89. The Autonomous Mobile Robots and Its Applications. IROS '89. Proceedings., IEEE/RSJ International Workshop on*, pages 12–23, Sep 1989.
- [48] Berthold K.P. Horn and Brian G. Schunck. Determining optical flow. Technical report, Cambridge, MA, USA, 1980.
- [49] S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Comput. Surv.*, 27(3):433–466, 1995.
- [50] Olivier D. Faugeras, editor. *Computer Vision - ECCV'90, First European Conference on Computer Vision, Antibes, France, April 23-27, 1990, Proceedings*, volume 427 of *Lecture Notes in Computer Science*. Springer, 1990.
- [51] Parag H. Batavia, Dean A. Pomerleau, and Charles E. Thorpe. Overtaking vehicle detection using implicit optical flow. In *Proceedings of the IEEE Intelligent Transportation Systems Conference*, pages 729–734, 1997.
- [52] J.M. Collado, C. Hilario, A. de la Escalera, and J.M. Armingol. Model based vehicle detection for intelligent vehicles. In *Intelligent Vehicles Symposium, 2004 IEEE*, pages 572–577, June 2004.
- [53] Thomas Zielke, Michael Brauckmann, and Werner von Seelen. Intensity and edge-based symmetry detection with an application to car-following. *CVGIP: Image Underst.*, 58(2):177–190, 1993.
- [54] Zehang Sun, G. Bebis, and R. Miller. On-road vehicle detection: a review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(5):694–711, May 2006.
- [55] U. Regensburger and V. Graefe. Visual recognition of obstacles on roads. In *Intelligent Robots and Systems '94. 'Advanced Robotic Systems and the Real World', IROS '94. Proceedings of the IEEE/RSJ/GI International Conference on*, volume 2, pages 980–987 vol.2, Sep 1994.
- [56] Massimo Bertozzi, Alberto Broggi, and Stefano Castelluccio. A real-time oriented system for vehicle detection. *J. Syst. Archit.*, 43(1-5):317–325, 1997.
- [57] L. Matthews, T. Ishikawa, and S. Baker. The template update problem. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):810–815, June 2004.
- [58] D. Pomerleau. Ralph: rapidly adapting lateral position handler. In *Intelligent Vehicles '95 Symposium., Proceedings of the*, pages 506–511, Sep 1995.
- [59] J.D. Crisman and C.E. Thorpe. Unscarf-a color vision system for the detection of unstructured roads. In *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*, pages 2496–2501 vol.3, Apr 1991.
- [60] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons Inc, 1973.

- [61] M. Bertozzi, A. Broggi, G. Conte, and A. Fascioli. Obstacle and lane detection on argo. In *Intelligent Transportation System, 1997. ITSC '97., IEEE Conference on*, pages 1010–1015, Nov 1997.
- [62] A. Ali and S. Afghani. Shadow based on-road vehicle detection and verification using haar wavelet packet transform. In *Information and Communication Technologies, 2005. ICICT 2005. First International Conference on*, pages 346–346, Aug. 2005.
- [63] João Manuel Leite da Silva. Fusão sensorial e comportamentos na equipa de futebol robótico cambada. Master’s thesis, Universidade de Aveiro, 2008.
- [64] A. Broggi, M. Bertozzi, A. Fascioli, C. Lo, and B. Piazzzi. The argo autonomous vehicle’s vision and control systems, 1999.
- [65] J. R. Parker. *Algorithms for Image Processing and Computer Vision*. John Wiley & Sons, Inc., New York, NY, USA, 1996.
- [66] MS Nixon and AS Aguado. *Feature Extraction & Image Processing*. Butterworth Heinmann/ Newnes, January 2002.
- [67] Chu Jiangwei, Ji Lisheng, Guo Lie, Libibing, and Wang Rongben. Study on method of detecting preceding vehicle based on monocular camera. In *Intelligent Vehicles Symposium, 2004 IEEE*, pages 750–755, June 2004.
- [68] Kah Phooi SENG King Hann LIM, Li Minn ANG and Siew Wen CHIN. Lane-vehicle detection and tracking. In *Proceedings of the International MultiConference of Engineers and Computer Scientists 2009*, volume 2, March 2009.
- [69] Wei Liu, Chunyan Song, XueZhi Wen, Huai Yuan, and Hong Zhao. A monocular-vision rear vehicle detection algorithm. In *Vehicular Electronics and Safety, 2007. ICVES. IEEE International Conference on*, pages 1–6, Dec. 2007.
- [70] Greg Welch and Gary Bishop. An introduction to the kalman filter.
- [71] A.S. Paul and E.A. Wan. Dual kalman filters for autonomous terrain aided navigation in unknown environments. In *Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on*, volume 5, pages 2784–2789 vol. 5, July-4 Aug. 2005.